

## CONNECTING CO INTENSITY MAPPING TO MOLECULAR GAS AND STAR FORMATION IN THE EPOCH OF GALAXY ASSEMBLY

TONY Y. LI<sup>1</sup>, RISA H. WECHSLER<sup>1,2</sup>, KIRUTHIKA DEVARAJ<sup>1</sup>, SARAH E. CHURCH<sup>1</sup>

<sup>1</sup>Kavli Institute for Particle Astrophysics and Cosmology;  
Physics Department, Stanford University, Stanford, CA 94305, USA;  
tonyyli@stanford.edu, rwechsler@stanford.edu and

<sup>2</sup> SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

(Dated: February 3, 2016)  
*Draft version February 3, 2016*

### ABSTRACT

Intensity mapping, which images a single spectral line from unresolved galaxies across cosmological volumes, is a promising technique for probing the early universe. Here we present predictions for the intensity map and power spectrum of the CO(1-0) line from galaxies at  $z \sim 2.4\text{--}2.8$ , based on a parameterized model for the galaxy-halo connection, and demonstrate the extent to which properties of high-redshift galaxies can be directly inferred from such observations. We find that our fiducial prediction should be detectable by a realistic experiment. Motivated by significant modeling uncertainties, we demonstrate the effect on the power spectrum of varying each parameter in our model. Using simulated observations, we infer constraints on our model parameter space with an MCMC procedure, and show corresponding constraints on the  $L_{\text{IR}}\text{--}L_{\text{CO}}$  relation and the CO luminosity function. These constraints would be complementary to current high-redshift galaxy observations, which can detect the brightest galaxies but not complete samples from the faint end of the luminosity function. By probing these populations in aggregate, CO intensity mapping could be a valuable tool for probing molecular gas and its relation to star formation in high-redshift galaxies.

*Subject headings:* galaxies: high-redshift — galaxies: evolution — ISM: molecules

### 1. INTRODUCTION

Current progress in cosmology and galaxy formation is strongly informed by galaxy observations at increasingly high redshifts, reaching into the epoch of reionization. Modern observatories such as ALMA may have even begun to probe the interstellar medium (ISM) of “typical” high-redshift galaxies (e.g. Riechers et al. 2014; Watson et al. 2015). Despite this, it remains difficult and expensive to observe populations of such galaxies in statistically large and complete samples, due to their faintness.

However, a full understanding of both cosmology and galaxy formation depends on such samples. To inform galaxy formation models, individual galaxies need to be placed in the context of their broader populations, for without such context it is unclear how representative individually observed galaxies may be. For cosmology, precise clustering measurements that trace large-scale structure require sufficient number densities of spectroscopically observed galaxies (Blake et al. 2011; Anderson et al. 2012).

*Intensity mapping* is a technique that can potentially address this challenge in regimes that are inaccessible by typical galaxy surveys. The method aims to map a single spectral line across cosmologically large volumes. The angular resolution used to do this images cumulative emission from multiple galaxies, instead of resolving single galaxies. In this way, intensity mapping aggregates flux from galaxies that would individually be below the detection limit, while still resolving large-scale structure by measuring intensity fluctuations over cosmological distances. As a consequence, this technique will necessarily describe galaxies in a statistical sense through measures like the spatial power spectrum.

In this paper, we look at intensity mapping of carbon monoxide (CO), a common tracer of molecular gas and star formation in nearby galaxies. After H<sub>2</sub>, CO is the most

abundant molecular species, tracing the metal-enriched, relatively dense ( $\gtrsim 10^2 \text{ cm}^{-3}$ ), cool to warm molecular ISM phase where stars form efficiently. This physically motivates the empirical conversion between CO and H<sub>2</sub> (for a review, see Bolatto et al. 2013), and by extension star formation. The CO lines themselves arise from a “ladder” of rotational transitions starting at millimeter wavelengths. For this study, we focus on the ground-state CO(1-0) transition at 115.27 GHz (2.6 mm).

Other lines besides CO have also been proposed as candidate lines for intensity mapping. The H I 21 cm line, in particular, has been extensively studied for the purposes of tracing large-scale structure out to  $z \sim 2.5$  as well as imaging hydrogen reionization at  $z \gtrsim 6$  (Chang et al. 2008; Morales & Wyithe 2010; Bandura et al. 2014). Additional potential targets include [CII] at 158  $\mu\text{m}$  (Gong et al. 2012; Silva et al. 2015; Uzgil et al. 2014), Ly $\alpha$  at 1216  $\text{\AA}$  (Pullen et al. 2014), and various other fine structure lines (Visbal & Loeb 2010), each with its own advantages and challenges.

Observing CO—as well as H I, [CII], and other lines—is necessary for a census of all phases of the ISM in galaxies. However, a practical advantage of CO is that it simultaneously emits at multiple frequencies: because of the small excitation energies of the lowest energy levels (5.5 and 16.6 K for the  $J = 1$  and 2 levels, respectively), a galaxy is likely to have comparable emission in CO(1-0) and CO(2-1). Observed signals in carefully chosen frequency bands can potentially be cross-correlated to determine the contribution of the CO-only signal.

Our focus is the “epoch of galaxy assembly,” at redshifts of roughly  $z \sim 2\text{--}3$ . This is a particularly interesting epoch for galaxy formation, because it is near the peak of cosmic star formation. Current and near-future galaxy surveys are also expected to reach into these redshifts, allowing the opportunity to cross-correlate the CO signal with galaxy surveys (Pullen

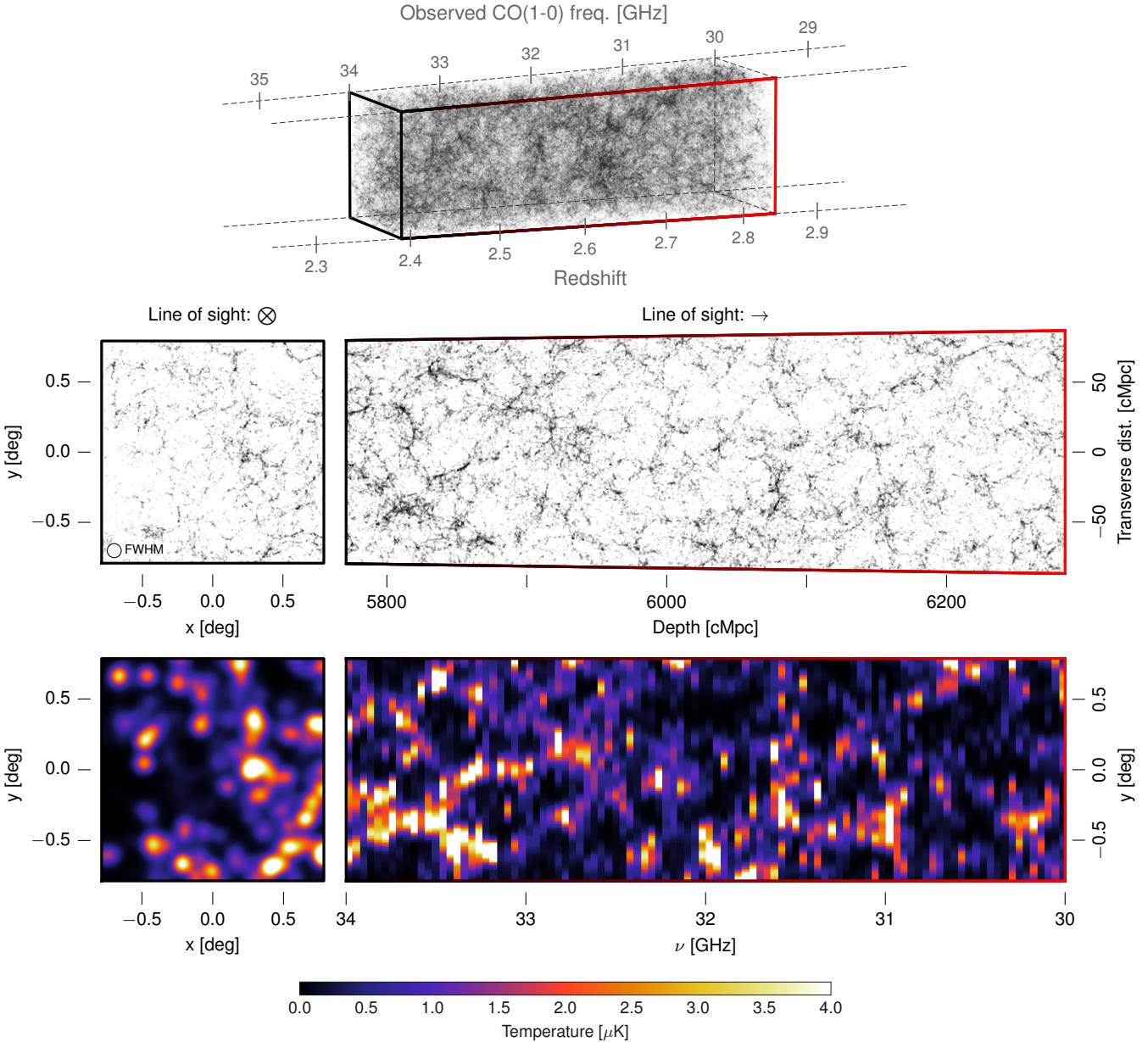


FIG. 1.— Input and output of our modeling process, i.e. initial dark matter halos and final CO intensity map (details in §2.2). These plots illustrate one realization of the pathfinder experiment’s survey volume (§2.4 and Table 2), while the full experiment’s survey area is 2.5 times larger. **Top:** Halos in the 3D volume, rendered to scale in comoving distance. Along the line-of-sight direction, we label the equivalent cosmological redshifts and redshifted CO(1-0) frequencies. **Middle:** 2D projections of halo positions. The left image shows the “front” view of halos that would fall into the highest 40 MHz frequency channel, or lowest redshift slice. The pathfinder beam size is shown for scale. The right image shows the “side” view of halos to a depth of 6 arcmin, or one beam width. **Bottom:** CO intensity map produced by our fiducial model. The slice volumes are the same as above, albeit with comoving depth converted to observed frequency. The same large-scale structure is readily apparent in both images, even with the lower resolution of the intensity map. The analysis in this paper relies on the power spectrum of this map (see Fig. 3).

et al. 2013). Our current understanding of star formation and gas content in this epoch is incomplete, and largely limited to the bright end of the relevant populations. In the longer term, observations at these redshifts could serve as a stepping stone for future CO observations that reach into the epoch of reionization (Carilli 2011; Gong et al. 2011; Lidz et al. 2011).

Previous predictions for the intensity of the CO signal vary by more than an order of magnitude (Breyssse et al. 2014, at  $z \sim 3$ ). The wide range simply reflects the current scarcity of data for typical high-redshift galaxies. It is possible to directly simulate these galaxies, but such simulations are expensive and still are quite uncertain. These uncertainties suggest

a need for alternative probes of high-redshift galaxy populations, especially over numbers and/or volumes currently inaccessible to traditional surveys.

Given the modeling uncertainties, predictions of the expected signal will only go so far, at least until a measurement is attempted. Here we also ask, what could we *learn* from intensity mapping if a measurement is made? More precisely: given hypothetical but tractable intensity mapping observations, what can we infer about the properties and distribution of the underlying galaxy population? To our knowledge, this question has not yet been directly addressed in the literature. Here we put these questions in the context of CO surveys that

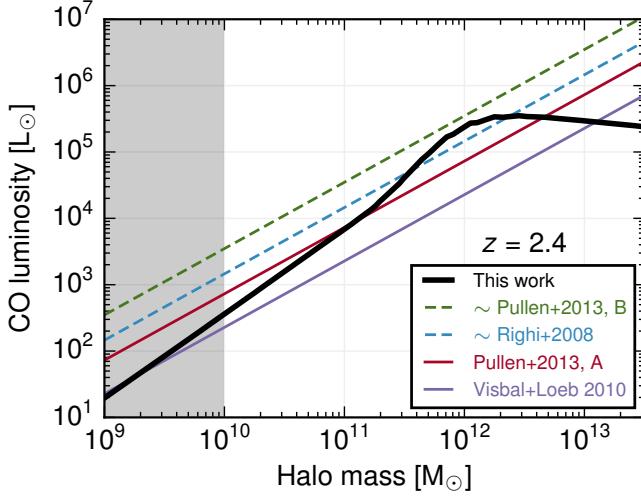


FIG. 2.— Relation between mean CO luminosity and halo mass at  $z = 2.4$ , at the low end of the redshift range in this study. The black solid line shows the relation used in this work. For comparison,  $L_{\text{CO}}(M_h)$  from previous studies (Righi et al. 2008; Visbal & Loeb 2010; Pullen et al. 2013) have also been plotted<sup>a</sup>. Solid color lines indicate the exact linear  $L_{\text{CO}}(M_h)$  relations used in those studies. Dotted color lines indicate approximate linear scalings for models without a direct  $L_{\text{CO}}(M_h)$  relation (as in Breysse et al. 2014). In our model, we imposed a minimum CO-luminous halo mass of  $10^{10} M_\odot$ , excluding halos in the shaded gray range (see Appendix A for some justification).

<sup>a</sup> To compare models consistently, the  $L_{\text{CO}}(M_h)$  relations plotted here have absorbed a “duty-cycle” factor, calculated in those papers as  $f_{\text{duty}} \approx 10^8 \text{ yr}/t_{\text{age}}(z)$ . Carilli (2011), Gong et al. (2011), and Lidz et al. (2011) are omitted from this plot because they focus only on reionization redshifts ( $z \gtrsim 6$ ).

are now being built or planned for the near future.

The structure of this paper is as follows. In §2, we summarize our method for (1) modeling galaxy CO luminosities and generating CO intensity maps and power spectra and (2) inferring constraints on this model given a hypothetical observation. In §3, we show how the variation of different model parameters affects our fiducial prediction. We then present MCMC-inferred parameter constraints from mock intensity mapping observations, as well as the inferred  $L_{\text{IR}}-L_{\text{CO}}$  relation and CO luminosity function. In §4, we discuss our results, the uncertainties in our modeling, and the ability of intensity mapping to constrain the properties of galaxy populations. Finally, in §5, we resummarize our main results and conclude.

## 2. METHODS

### 2.1. Context: Previous Studies

Several previous studies have explored CO intensity mapping and modeled the expected signal. To date, they include: Righi et al. (2008), Visbal & Loeb (2010), Gong et al. (2011), Carilli (2011), Lidz et al. (2011), Pullen et al. (2013).

Since the models in those studies inform our own, the reader may find a brief summary useful. Table 3, located toward the end of this paper, compares the most relevant aspects of each model. The common thread across nearly all of them (except Gong et al. 2011) is an assumed connection between star-formation rate (SFR) and CO luminosity. In those papers, the procedure is to first calculate SFRs and then convert to CO luminosity, though the details vary.

Around the redshifts considered in this study, Breysse et al. (2014) found that the mean CO intensity predicted by some of these models differs at the extremes by more than an order of magnitude, which simply reflects the significant uncertainty

in these predictions.

### 2.2. Part 1: Modeling CO luminosity

In our modeling, we begin with dark matter halos (tracing underlying cosmological structure) and ultimately generate simulated three-dimensional CO intensity maps.

In order, the components of this calculation are:

1. Dark matter halos (§2.2.1)
2. Star-formation rates (§2.2.2)
3. Infrared luminosities (§2.2.3)
4. CO luminosities (§2.2.4)
5. Intensity maps and power spectra (§2.2.5)

Figure 1 illustrates the input (dark matter halos) and output (CO intensity map) of this process. It also provides some visual intuition for the dimensions of the survey volume, as well as the cosmological structure contained within.

We discuss each of the individual steps in more detail below. Where parameters are introduced, we state their fiducial values, as well as the priors adopted for the MCMC analysis (§2.3).

#### 2.2.1. Dark Matter Halos

We begin with dark matter halos in a “lightcone” volume. To obtain this, we use the results of a cosmological  $N$ -body dark matter simulation (the c400–2048 box<sup>1</sup>). The simulation was run with L-GADGET (based on GADGET-2, Springel et al. 2001; Springel 2005). The box has 2048<sup>3</sup> particles and a side length of 400 Mpc  $h^{-1}$ , resulting in a particle mass of  $5.9 \times 10^8 M_\odot h^{-1}$ . The softening length used is 5.5 kpc  $h^{-1}$ , constant in comoving length. The initial conditions are generated by 2LPTIC<sup>2</sup> (Crocce et al. 2006) at  $z = 99$ , with the power spectrum generated by CAMB<sup>3</sup>.

The  $\Lambda$ CDM cosmological parameters of the simulation are  $\Omega_m = 0.286$ ,  $\Omega_\Lambda = 0.714$ ,  $\Omega_b = 0.047$ ,  $h = 0.7$ ,  $\sigma_8 = 0.82$ , and  $n_s = 0.96$ . We assume these values throughout this paper.

Dark matter halos were identified at each simulation snapshot using the ROCKSTAR halo finder (Behroozi et al. 2013c). In this work, we treated subhalos and central halos identically, but excluding subhalos from our calculations does not significantly affect our results.

The simulation output consisted of 100 snapshots, from  $z = 12.33$  to  $z = 0$  inclusive, with equal logarithmic spacing in  $(1+z)^{-1}$ , though only the halos from six of these snapshots were ultimately used. Lightcones were generated from these snapshots by choosing an arbitrary  $z = 0$  observer origin and direction, then selecting all halos along the line of sight within the desired survey area and redshift range (see §2.4) from the appropriate redshift snapshots.

#### 2.2.2. Halos and Star Formation

To get the SFR of a dark matter halo, we use the results of Behroozi et al. (2013a) and Behroozi et al. (2013b), which empirically quantified the stellar mass history of dark matter

<sup>1</sup> Provided by Matthew Becker (M. Becker et al. 2015, in preparation)

<sup>2</sup> <http://cosmo.nyu.edu/roman/2LPT/>

<sup>3</sup> <http://camb.info/>

TABLE 1  
MODEL PARAMETERS

Label	Link	Description	Fiducial	Prior	Details
$\sigma_{\text{SFR}}$	Halos $\rightarrow$ SFR	Log-scatter in SFR	0.3	$\mathcal{N}(0.3, 0.1)$ and $\sigma_{\text{SFR}} \geq 0$	§2.2.2
$\log \delta_{\text{MF}}$	SFR $\rightarrow L_{\text{IR}}$	SFR- $L_{\text{IR}}$ scaling	0.0	$\mathcal{N}(0.0, 0.3)$	§2.2.3, Eq. 1
$\alpha$	$L_{\text{IR}} \rightarrow L_{\text{CO}}$	$L_{\text{IR}}\text{-}L_{\text{CO}}$ log-slope	+1.37	$\mathcal{N}(1.17, 0.37)$	§2.2.4, Eq. 2
$\beta$	$L_{\text{IR}} \rightarrow L_{\text{CO}}$	$L_{\text{IR}}\text{-}L_{\text{CO}}$ log-intercept	-1.74	$\mathcal{N}(0.21, 3.74)$	§2.2.4, Eq. 2
$\sigma_{L_{\text{CO}}}$	$L_{\text{IR}} \rightarrow L_{\text{CO}}$	Log-scatter in $L_{\text{CO}}$	0.3	$\mathcal{N}(0.3, 0.1)$ and $\sigma_{L_{\text{CO}}} \geq 0$	§2.2.4

NOTE. —  $\mathcal{N}(\mu, \sigma)$  denotes a normal distribution with mean  $\mu$  and variance  $\sigma^2$ .

TABLE 2  
INSTRUMENT AND SURVEY PARAMETERS

Type	Description	Pathfinder	Full
Instrument	System temperature ( $T_{\text{sys}}$ )	40 K	35 K
	Dual-polarization feeds ( $N_{\text{feeds}}$ ) <sup>a</sup>	19	500
	Beam width ( $\theta_{\text{FWHM}}$ )	6'	3'
	Frequency band ( $\Delta\nu$ )	30-34 GHz	30-34 GHz
	Frequency channels ( $\delta_{\nu}$ )	40 MHz	10 MHz
Survey	Survey area / patch ( $\Omega_{\text{surv}}$ )	2.5 deg <sup>2</sup>	6.25 deg <sup>2</sup>
	On-sky time / patch ( $\tau_{\text{obs}}$ ) <sup>b</sup>	1500 hr	2250 hr
Derived	CO(1-0) redshift range	2.4–2.8	2.4–2.8
	Instrument sensitivity	1026 $\mu\text{K s}^{1/2}$	783 $\mu\text{K s}^{1/2}$
	Final map sensitivity <sup>c</sup>	41.5 $\mu\text{K MHz}^{1/2}$	18.3 $\mu\text{K MHz}^{1/2}$
	Total $P(k)$ detection significance ( $\text{SNR}_{\text{tot}}$ ) <sup>d</sup>	7.89 $\sigma$	144 $\sigma$

<sup>a</sup>Dual polarization effectively doubles the number of feeds, so in our calculations we model  $N_{\text{feeds}} = 19$  (500) as 38 (1000) feeds.

<sup>b</sup>We assume that observing time is divided equally over four patches with  $\sim 35\%$  observing efficiency. Thus, 750 hr per patch equals a total of  $\sim 1$  year of operation.

<sup>c</sup>This sensitivity measure is proportional to the 3D noise power spectral density (Eq. C2). It is independent of spectral resolution, which is important here since we are mainly interested in measuring 3D structure. To get rms noise per channel instead, divide the value by  $\sqrt{\delta_{\nu}}$ .

<sup>d</sup>See §3.1 for additional discussion.

halos back to  $z = 8$ . Briefly, that study constrained a parameterized stellar mass–halo mass relation by applying it to simulated halo merger trees, accounting for systematic errors and biases, and comparing derived stellar mass functions, cosmic SFRs, and specific SFRs with a comprehensive compilation of observational data. The reader is referred to the original papers for a more detailed discussion of that study.

For our purposes, we are primarily interested in  $\overline{\text{SFR}}(M, z)$ , their derived results for the *average* halo SFR for a given halo mass and redshift. To simplify this calculation, we interpolate their tabulated data<sup>4</sup> for  $\overline{\text{SFR}}(M, z)$ .

The results of Behroozi et al. (2013a) also provide  $\pm 1\sigma$  posterior constraints on  $\overline{\text{SFR}}$ . Note that this is distinct from halo-to-halo scatter, which we address later. However, because these bounds span a fairly consistent log-space interval ( $\sim \pm 0.15$  dex) over the relevant masses and redshifts, any variation within them is effectively a rescaling factor, which we absorb into the parameter  $\delta_{\text{MF}}$  (see below, §2.2.3, Eq. 1).

We also add halo-to-halo log-normal scatter, parameterized as  $\sigma_{\text{SFR}}$ , since the results of Behroozi et al. (2013a) constrain only average halo SFRs. This single-parameter scatter is a simple way of capturing the variation in SFR for a given halo mass. There is evidence that “normal” star-forming galaxies exhibit a strong correlation with their stellar mass, with a scatter of  $\sim 0.2$ –0.4 dex, while starbursts may exist as a separate population with unusually high SFR (e.g. Speagle et al. 2014; Salmon et al. 2015). We assume the scatter in SFR given stellar mass to be reasonably similar to the scatter in SFR given halo mass. With this in mind, we choose a fiducial value of  $\sigma_{\text{SFR}} = 0.3$  and a prior of  $\sigma_{\text{SFR}} = 0.3 \pm 0.1$ . To avoid unphysical

negative scatter, we also require  $\sigma_{\text{SFR}} \geq 0$ .

### 2.2.3. Star Formation and Infrared Luminosity

Empirically connecting SFR to  $L_{\text{CO}}$  requires at least one intermediate, directly observable quantity, since SFRs are not directly measured. Instead, they are inferred from photometric or spectral tracers with various underlying assumptions. We will take this intermediate tracer to be the total infrared luminosity  $L_{\text{IR}}$  (see Carilli 2011; Lidz et al. 2011; Pullen et al. 2013), conventionally the integrated 8–1000  $\mu\text{m}$  luminosity. Thus, the model is calibrated on empirical correlations between SFR,  $L_{\text{IR}}$ , and  $L_{\text{CO}}$ , resting on physical assumptions about star formation, dust, and molecular gas. Some recent work has focused on quantifying these relations at higher redshifts (e.g. Béthermin et al. 2015, for  $z \lesssim 4$ ), but significant uncertainties remain, especially in fainter galaxies and at higher redshifts, and part of our aim is to analyze how the signal may vary within those uncertainties.

We assume a correlation between SFR and  $L_{\text{IR}}$  (Kennicutt 1998) of the form

$$\text{SFR} = \delta_{\text{MF}} \times 10^{-10} L_{\text{IR}} \quad (1)$$

where SFR is in units of  $\text{M}_{\odot} \text{yr}^{-1}$  and  $L_{\text{IR}}$  is in units of  $\text{L}_{\odot}$ . The normalization  $\delta_{\text{MF}}$  is sensitive to assumptions about the initial mass function, the duration of star formation, and dust, discussed in more detail in §4.2. However, it is generally calculated to be a factor of order unity ( $0.8 \lesssim \delta_{\text{MF}} \lesssim 2$  in, e.g. Scoville & Young 1983; Thronson & Telesco 1986; Kennicutt 1998; Barger et al. 2000; Rowan-Robinson 2000; Omont et al. 2001).

We adopt a Chabrier (2003) initial mass function, which conventionally entails  $\delta_{\text{MF}} = 1.0$  (e.g. Magnelli et al. 2012;

<sup>4</sup> Available at the time of writing at <http://www.peterbehroozi.com/data.html>

Carilli & Walter 2013), as was used by Behroozi et al. (2013a).

To our knowledge, there has not been a recent comprehensive study of the expected range of  $\delta_{\text{MF}}$  for high-redshift galaxies, and such a study is beyond the scope of this paper. In order to remain consistent with the values quoted above, we adopt a log-normal prior of  $\log \delta_{\text{MF}} = 0.0 \pm 0.3$  ( $\delta_{\text{MF}} \approx 1.0^{+2.0}_{-0.5}$ ). Note that this means the prior's  $\pm 3\sigma$  interval spans nearly 2 dex.

Eq. 1 could also be written as  $SFR = \delta_{\text{MF}} \delta_{\text{SB}} \times 10^{-10} L_{\text{IR}}$ , where  $\delta_{\text{SB}}$  explicitly accounts for the fraction of  $L_{\text{IR}}$  due to starbursts, as opposed to active galactic nuclei or other sources. For simplicity, we have absorbed  $\delta_{\text{SB}}$  into  $\delta_{\text{MF}}$ , and any scatter about Eq. 1 due to variation in galaxy type is assumed to be absorbed into  $\sigma_{\text{SFR}}$  (§2.2.2).

#### 2.2.4. Infrared Luminosity and CO Luminosity

To convert infrared luminosity to CO luminosity, we assume a power-law relation of the form

$$\log L_{\text{IR}} = \alpha \log L'_{\text{CO}} + \beta \quad (2)$$

where  $L_{\text{IR}}$  is in units of  $L_{\odot}$ , and  $L'_{\text{CO}}$  is in units of  $\text{K km s}^{-1} \text{ pc}^2$  (observer units for velocity- and area-integrated brightness temperature).

For our fiducial model, we use the fit from Carilli & Walter (2013), which found from a census of high-redshift galaxies  $\alpha = 1.37(\pm 0.04)$  and  $\beta = -1.74(\pm 0.40)$ .

More generally, however, this relation is a rough proxy for the relation between star formation and molecular gas, and has been fit to data from high-redshift galaxies ( $z \gtrsim 1$ , but typically few if any galaxies beyond  $z \sim 3$ ) in a number of studies. These studies have found, e.g.,  $(\alpha, \beta)$  values of  $(1.13, 0.53)$ ,  $(1.37, -1.74)$ ,  $(1.00, 2.00)$ ,  $(1.17, 0.28)$  (Daddi et al. 2010; Carilli & Walter 2013; Greve et al. 2014; Dessauges-Zavadsky et al. 2015, respectively). These values are closely fit by the line

$$\alpha \approx 0.10\beta + 1.19 \quad (3)$$

as they are generally derived from similar samples of the most luminous, and therefore detectable, high- $z$  galaxies. We choose deliberately loose priors of  $\alpha = 1.17 \pm 0.37$  and  $\beta = 0.21 \pm 3.74$  by taking the mean of those four values and using their full range as the  $1\sigma$  spread.<sup>5</sup>

We also add log-normal scatter to  $L'_{\text{CO}}$ , fiducially parameterized by  $\sigma_{L'_{\text{CO}}} = 0.3$ . This is effectively the scatter in  $L'_{\text{CO}}$  for a given  $L_{\text{IR}}$  or SFR (they scale linearly and so are separated in log-space by a constant). However, equal values of  $\sigma_{L'_{\text{CO}}}$  and  $\sigma_{\text{SFR}}$  do not necessarily have the same effect on the signal, since SFR and  $L'_{\text{CO}}$  are not required to scale linearly. In principle, though,  $\sigma_{L'_{\text{CO}}}$  and  $\sigma_{\text{SFR}}$  could be combined into a single scatter parameter. We choose not to do so because empirical constraints, which inform our priors, are generally on the two separate parameters.

The prior on this parameter is  $\sigma_{L'_{\text{CO}}} = 0.3 \pm 0.1$ , with the requirement that  $\sigma_{L'_{\text{CO}}} \geq 0$ . This is consistent with the results of aforementioned studies, where a scatter has been quoted.

Note that the conversion from  $L'_{\text{CO}}$  (units of  $\text{K km s}^{-1} \text{ pc}^2$ )

<sup>5</sup> While this paper was under review, the values of  $(\alpha, \beta)$  in Dessauges-Zavadsky et al. (2015) were updated from  $(1.19, 0.05)$  to  $(1.17, 0.28)$ . We have updated text and figures in this work where appropriate.

to  $L_{\text{CO}}$  (units of  $L_{\odot}$ ) is

$$L_{\text{CO}} = 4.9 \times 10^{-5} L_{\odot} \left( \frac{\nu_{\text{CO, rest}}}{115.27 \text{ GHz}} \right)^3 \left( \frac{L'_{\text{CO}}}{\text{K km s}^{-1} \text{ pc}^2} \right) \quad (4)$$

where  $\nu_{\text{CO, rest}} = 115.27$  GHz is the rest-frame frequency of the CO transition.

To resummarize the model:

1. Halos  $\rightarrow$   $SFR$ : Get  $\overline{SFR}(M, z)$  from the results of Behroozi et al. (2013a)
2. Add log-scatter,  $\sigma_{\text{SFR}}$
3.  $SFR \rightarrow L_{\text{IR}}$ : Get  $L_{\text{IR}}$  from  $SFR = \delta_{\text{MF}} \times 10^{-10} L_{\text{IR}}$
4.  $L_{\text{IR}} \rightarrow L'_{\text{CO}}$ : Get  $L'_{\text{CO}}$  from  $\log L_{\text{IR}} = \alpha \log L'_{\text{CO}} + \beta$
5. Add log-scatter,  $\sigma_{L'_{\text{CO}}}$

with fiducial parameter values:

$$\sigma_{\text{SFR}} = 0.3, \sigma_{L'_{\text{CO}}} = 0.3, \\ \delta_{\text{MF}} = 1.0, \alpha = 1.37, \beta = -1.74$$

Figure 2 shows the combined result of these steps, plotting the mean  $L_{\text{CO}}(M_h)$  relation from our fiducial model, as well as the equivalent relation from previous studies. Notably,  $L_{\text{CO}}$  in this model is not linear in  $M$ , a simplifying assumption that has been adopted in some previous studies. This is a consequence of our choice to model the average population including quiescent galaxies, which results in a flatter function of mass at the high-mass end than a model constrained using only detected star-forming galaxies. This can affect the shape of the power spectrum, even if the mean CO brightness temperature is the same.

Our model may be seen as combining the two models from Pullen et al. (2013), “A” and “B”, in that we still start with dark matter halos (their model A) but calculate empirically constrained SFRs (their model B). In short, we combine the machinery for A with the motivation for B.

#### 2.2.5. Generating Intensity Maps and Power Spectra

Once halo CO luminosities are calculated, we generate 3D intensity maps (spectral data cubes) and power spectra. We calculate intensities as brightness temperatures, since we expect observation and calibration to be in terms of that quantity.

To do this, the halos are binned by their positions on an RA  $\times$  Dec  $\times$   $\nu_{\text{obs}}$  grid, where  $\nu_{\text{obs}} = \nu_{\text{CO, rest}} / (1+z)$ . From the total CO luminosity in each cell, brightness temperatures are calculated using the Rayleigh-Jeans relation

$$T = \frac{c^2 I_{\nu, \text{obs}}}{2k_B \nu_{\text{obs}}^2}. \quad (5)$$

Here,  $I_{\nu, \text{obs}} = L_{\text{CO}} / 4\pi D_L^2 \delta_{\nu}$ , where  $D_L$  is the luminosity distance to the source, and  $\delta_{\nu}$  is the width of the frequency channel (spectral resolution element), specified in §2.4. This assumes that the CO line profile is approximately a delta function. In this study, we have not included the effects of Doppler broadening and redshift-space distortions, which would somewhat alter the line-of-sight signal, but they are left for future studies.

We calculate power spectra  $P(\mathbf{k})$  by converting the grid to 3D comoving units and squaring the discrete Fourier transform of the temperature cube. More details are provided in

Appendix B. Averaging the 3D power spectrum  $P(\mathbf{k})$  in radial bins gives us the spherically averaged power spectrum,  $P(k)$ . In this paper, we plot on the quantity  $\Delta^2(k) = k^3 P(k)/2\pi^2$  instead of  $P(k)$  directly, mainly to compare with previous studies. The physical meaning of  $\Delta^2(k)$  is the amount of variance in  $T$  contributed per logarithmic interval in  $k$ .<sup>6</sup>

We assume white (Gaussian) instrumental noise in these simulations. As with all measurements, actual data will have additional non-ideal signal components, but we do not attempt to add these instrument-specific terms here. By using  $P(k)$  as calculated above, we implicitly assume that the CO power spectrum has been cleaned of instrument noise. In this case, we can imagine having subtracted a constant noise power spectrum  $P_n(k) \propto \sigma_n^2$  (Eq. C2) from the noisy spectrum, leaving only residual fluctuations from noise variance. Error bars on  $P(\mathbf{k})$  have been calculated to account for this noise variance, as well as resolution limits and sample variance of  $P(k)$ . We detail this calculation in Appendix C.

When showing predictions for the power spectrum, we average over 100 realizations of the same volume, where each realization is taken from a separately generated lightcone. This averaging smoothes out fluctuations in the power spectrum that are purely sample variance. While those fluctuations due to sample variance are important to consider, they distract from the average power spectrum predicted by the model, which is what we aim to show unless otherwise specified.

### 2.3. Part 2: MCMC Inference From Mock Signal

Modeling the expected signal is a useful first step. However, given the significant empirical uncertainties of this (or any) model, there is a limit to how useful a single prediction can be.

A question of at least equal importance is how intensity mapping can inform our understanding of the underlying galaxy population. Additionally, it is worth addressing the implications for galaxy properties *if no clear CO signal* is detected by the experiment: given the uncertainties, this is a possibility. While a non-detection might not be the most desired outcome, we wish to know whether it would be informative and therefore scientifically interesting. To that end, we produce the following three “mock” power spectra:

1.  $P(k)$  from our fiducial  $L_{\text{CO}}(M)$  model.
2.  $P(k) \approx 0$  within random fluctuations ( $\sigma_P \sim P_n/\sqrt{N_{\text{modes}}}$ , the second term of Eq. C1). This represents a signal consistent with a non-detection after subtracting a flat noise spectrum.
3.  $P(k)$  from the simulated data of Obreschkow et al. (2009b), which were generated with a model different from ours.

Combined with priors informed by existing observations, we use the intensity mapping power spectrum to infer constraints on our model parameter space for all three of these models.

We generate these three mock spectra and their error bars for both proposed instruments discussed in the following section (§2.4), for a total of six hypothetical signals. The purpose of the fiducial signal is to see how closely we can recover the “true” parameter values. The purpose of the “zero” power

<sup>6</sup>  $\Delta^2(k)$  is sometimes called the “dimensionless” power spectrum if the real-space quantity being analyzed is unitless, e.g. number overdensity. We avoid this label here because the real-space quantity of interest is brightness temperature, so  $\Delta^2(k)$  has units of  $\mu\text{K}^2$ .

spectrum is to investigate what constraints a non-detection places on the galaxy population. The purpose of the signal from Obreschkow et al. (2009b) data is to check whether our results are still reasonable when an independent model is used to generate the CO signal.

We use emcee<sup>7</sup>, an implementation of an affine-invariant ensemble Markov chain Monte Carlo (MCMC) algorithm (Foreman-Mackey et al. 2013; Goodman & Weare 2010), to sample the posterior distributions in our model parameter space. We provide more details, including the likelihood, in Appendix B.3.

A summary of the model parameters and their priors is provided in Table 1. All parameters are sampled in linear space except for  $\delta_{\text{MF}}$ , which is sampled in log space.

In this study we model  $\sigma_{\text{SFR}}$  and  $\sigma_{L_{\text{CO}}}$  as uncorrelated, so from the steps described in §2.2, one could write the total scatter as  $\sigma_{\text{tot}} = (\sigma_{\text{SFR}}^2/\alpha^2 + \sigma_{L_{\text{CO}}}^2)^{1/2}$ . However, if we consider perfectly correlated or anticorrelated scatter, the total scatter would instead be  $\sigma_{\text{tot}} = \sigma_{\text{SFR}}/\alpha + \sigma_{L_{\text{CO}}}$  or  $\sigma_{\text{tot}} = |\sigma_{\text{SFR}}/\alpha - \sigma_{L_{\text{CO}}}|$ . This would change the fiducial value ( $\sigma_{\text{tot}} \approx 0.37$ ) to  $\sigma_{\text{tot}} \approx 0.52$  or  $\sigma_{\text{tot}} \approx 0.08$ , respectively. Insofar as we only care about the total scatter, these extreme cases should be adequately covered by our priors. Strongly correlated scatter could plausibly affect the inferred values of other parameters, but we do not model this for the present study.

In this work, we have chosen to err on the side of broad priors. A more thorough analysis of the uncertainties, systematics, and details outside the scope of this study is warranted and welcome in future papers.

### 2.4. Instrument and Survey Design

To connect this work to proposed and feasible observations, we consider two examples of possible dedicated instruments for CO intensity mapping that can be built with current detector technology. In our calculations, the instrument and survey parameters determine the size and resolution of our 3D intensity map, as well as the error bars on the power spectrum.

An essential requirement of intensity mapping is high frequency resolution. In contrast to current measurements of extragalactic backgrounds—such as the cosmic infrared background (e.g. with *Spitzer* or *Herschel*) or the cosmic microwave background (CMB)—which characterize excess radiation in relatively broad bands, the redshifted CO intensity from galaxies will have narrow fluctuations in frequency space, reflecting cosmological structure along the line of sight. For reference, for CO(1-0) emission from  $z = 2.5$  ( $\nu_{\text{obs}} \approx 33$  GHz), a spatial depth of 1 comoving Mpc corresponds to a frequency width of 8.25 MHz, which requires high-resolution spectroscopy to resolve.

Beyond that, both instruments we consider are single-dish, ground-based telescopes covering a band of 30–34 GHz in the Ka microwave band. These frequencies correspond to redshifted CO(1-0) emission from  $z \approx 2.4$ –2.8.

The first instrument is envisioned as a pathfinder experiment, with parameters similar to the CO Mapping Array Pathfinder (COMAP), currently under development. Its main goal would be to detect or constrain the CO signal from galaxies at  $z \sim 2$ –3. We have assumed an aperture of  $\sim 6$  m with a frequency resolution of  $R = \nu_{\text{obs}}/\delta_\nu \approx 800$ , with a  $2.5 \text{ deg}^2$  field of view. For the rest of the paper, we refer to this instrument as the “pathfinder” experiment.

<sup>7</sup> <http://dan.iel.fm/emcee>

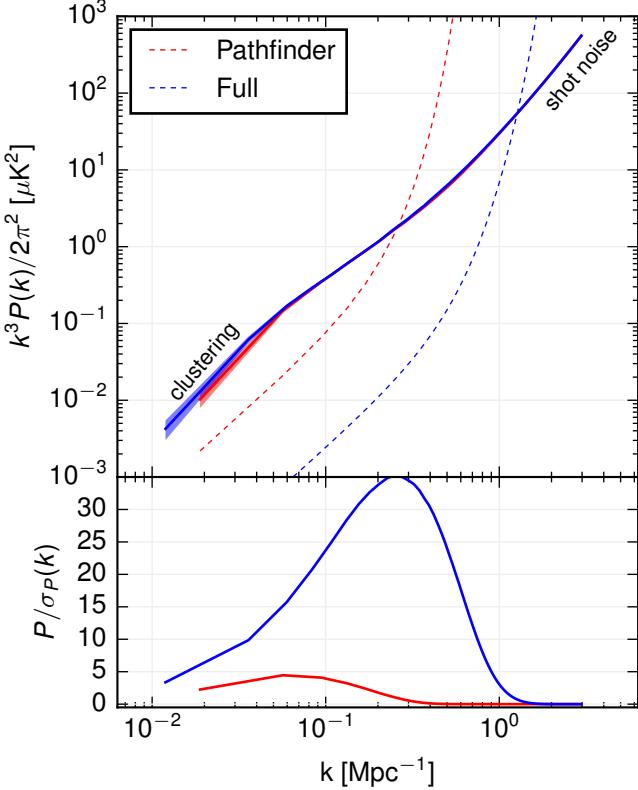


FIG. 3.— Fiducial CO power spectrum and detection significance. **Top:** Spherically averaged CO power spectrum from our fiducial model (§2.2), as measured by the pathfinder (red) and full (blue) experiments and averaged over four separate sky patches. *Solid lines* show the measured average power spectrum. *Shaded regions* show the  $1\sigma$  uncertainty from sample variance. *Dashed lines* indicate the  $1\sigma$  limits from thermal instrument noise. The upturn in these limits at high  $k$  (small scales) arises from resolution limits (finite beam and channel width). **Bottom:** Detection significance of the power spectrum as a function of  $k$ . The total detection significance (Eq. 7) is 7.89 (144) for the pathfinder (full) experiment.

The second instrument is a more advanced experiment, which would aim to measure the CO signal with greater sensitivity, over a wider area, and with better resolution. To this end, we have assumed an aperture of  $\sim 12$  m aperture with  $R \approx 3200$ , with a  $6.25 \text{ deg}^2$  field of view. For the rest of the paper, we refer to this instrument as the “full” experiment.

The relevant parameters from the two instruments are summarized in Table 2. For the purposes of our calculations, we assume that each instrument will pursue a survey that observes four separate patches on the sky, a necessary survey strategy due to the limitations of ground-based observations. However, we can combine the power spectra from each patch, reducing the error bars on the power spectrum (see Appendix C) by a factor of approximately  $\sqrt{4} = 2$ .

Throughout this paper, we use units of brightness temperature. If desired, the unit conversion from brightness temperature  $T$  to flux density  $S$  can be written as

$$S = 0.108 \text{ mJy beam}^{-1} \left( \frac{\nu_{\text{obs}}}{32 \text{ GHz}} \right)^2 \left( \frac{\theta_{\text{FWHM}}}{6'} \right)^2 \frac{T}{\mu\text{K}}. \quad (6)$$

#### 2.4.1. Sensitivity

We have given specific parameters for a CO intensity mapping experiment, but in this section we discuss, more broadly, the question of how we should characterize the sensitivity of

an intensity mapping experiment. There are at least a few distinct ways to address this:

- **Mean CO intensity vs instrument noise.** One can compare the mean CO brightness temperature  $\langle T_{\text{CO}} \rangle$  to the instrument noise fluctuations  $\sigma_n$  (Eq. C6), which is a useful first check (e.g. Carilli 2011, at  $z \gtrsim 6$ ). However, this is not exactly the signal being measured, and it is possible to have an instrument where  $\langle T_{\text{CO}} \rangle < \sigma_n$  in individual map “voxels” (as with the pathfinder experiment in this work), and still measure cosmological intensity fluctuations on larger scales. On those scales (across multiple voxels), random noise fluctuations should average out, while cosmological fluctuations should remain. Moreover, an actual observation would require subtraction of continuum foregrounds, making a direct measurement of  $\langle T_{\text{CO}} \rangle$  (the  $k = 0$  mode) difficult.
- **Measuring spatial fluctuations in the CO signal.** In fact, we are interested in detecting the *spatial structure* of the signal. Resolved 3D spatial structure is essential for intensity mapping. Without it, there is no way of determining whether a signal is CO from high-redshift galaxies, which trace cosmological structure along the line of sight, or simply an unwanted systematic (e.g. synchrotron foregrounds, which are generally smooth in frequency space). For an experiment with the goal of making a *detection*, we can sum in quadrature the detection significances of each measured mode (Pullen et al. 2013; Breysse et al. 2014, Eq. 7).

- **Usefulness for inferred constraints.** This requires a properly targeted range of  $k$ . To measure cosmological large-scale structure, e.g. baryon acoustic oscillations or the linear power spectrum, one would like to cover large areas, focusing mainly on the clustering modes at low  $k$ . To make inferences about the galaxy population, one also needs sensitivity to the shot noise modes at high  $k$ . However, the exact transition between the clustering and shot noise scales depends on the details of the underlying galaxy population, as shown in §3.2. One needs sensitivity to  $P(k)$  in both regimes, either from a single experiment or from multiple experiments targeting different scales. Information about the underlying galaxy population is contained in the relative strength of the power spectrum on different scales.

## 3. RESULTS

### 3.1. Fiducial Prediction and Detection Significance

Figure 3 shows the fiducial power spectra obtained from our mock intensity maps (e.g. Fig. 1), as well as the mode detection significance as a function of  $k$ . As has been noted by previous studies, the power spectrum may be thought of, in the context of the “halo model” formalism (Cooray & Sheth 2002), as having a clustering component (at low  $k$ ) and a shot noise component (at high  $k$ ).

In averaging  $P(k)$ , we choose bin widths of  $\Delta k = 2\pi/L_{\min}$ , where  $L_{\min}$  is the shortest comoving dimension of the survey volume, specifically  $\Delta k = 0.038 \text{ Mpc}^{-1}$  (for the pathfinder) and  $0.024 \text{ Mpc}^{-1}$  (for the full experiment). Note that the apparent deficit in the pathfinder’s low- $k$  power is because its wider, linearly spaced  $k$ -space bins cause the the average power in the lowest bin to appear underestimated.

Averaging the power spectrum over four identical patches, as mentioned in 2.4, we find that the maximum detection significance  $P(k)/\sigma_P(k)$  for any single  $k$  mode is 4.46 and 35.2

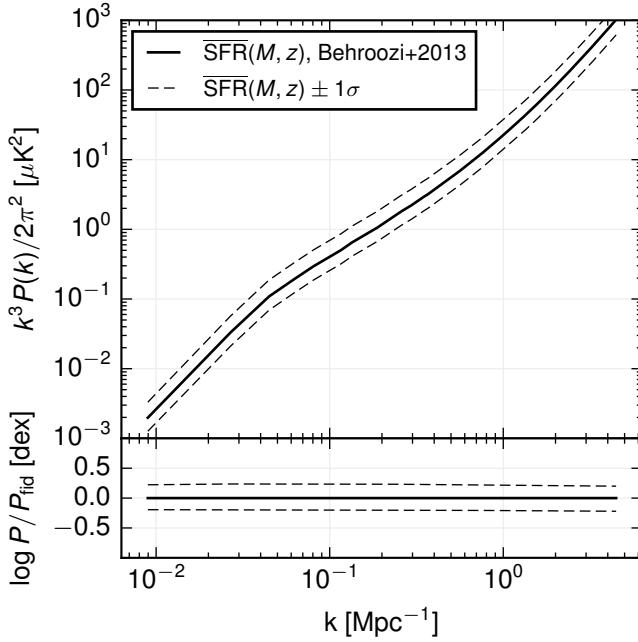


FIG. 4.— Range of power spectra spanned by  $\pm 1\sigma$  uncertainties on the mean  $SFR(M, z)$  from Behroozi et al. (2013a). At a fixed scale  $k$ , these span power spectra values of approximately  $\pm 0.2$  dex. No halo-to-halo scatter in  $L_{CO}$  has been included here (the effect of scatter is shown in Fig. 5).

for the pathfinder and full experiments, respectively. However, we can quantify the total detection significance of the power spectrum as (e.g. Pullen et al. 2013; Breysse et al. 2014)

$$\text{SNR}_{\text{tot}}^2 = \sum_i \left[ \frac{P(k_i)}{\sigma_P(k_i)} \right]^2 \quad (7)$$

where  $k_i$  are the discrete values of  $k$  at which the power spectrum is calculated, separated by  $\Delta k$ .

Assuming uncorrelated errors, we find the total signal-to-noise ratio (SNR) to be 7.89 and 144 for the pathfinder and full experiments, respectively. However, those values are only for our specific choice of modeling and observing parameters. In general,  $\text{SNR}_{\text{tot}}$  may be improved by increasing bandwidth or by decreasing noise fluctuations (the latter may be achieved by increasing observing time, increasing the number of feeds, or decreasing system temperature).

It is also possible to optimize the survey area to maximize  $\text{SNR}_{\text{tot}}$  (see, e.g. Breysse et al. 2014), though note that the exact optimal area depends on  $P(k)$  and is therefore model-dependent. When we perform this calculation, we obtain 0.6 and 9.2 deg<sup>2</sup> for the pathfinder and full experiments, respectively. These values differ from our choices of 2.5 and 6.25 deg<sup>2</sup>, but adopting the optimal values does not yield dramatic improvements (increasing  $\text{SNR}_{\text{tot}}$  to 9.48 and 146). Throughout this study, then, we maintain the pathfinder and full experiment parameters in Table 2. See Appendix D for more details and considerations.

### 3.2. Varying Model Parameters

The uncertainties in  $\overline{SFR}$  alone allow the power spectrum to span  $\pm 0.2$  dex in amplitude. This can be seen in Figure 4, which shows the range of power spectra expected from the the  $1\sigma$  posterior on  $SFR(M_h, z)$  (Behroozi et al. 2013a). As

noted, the width of this interval is fairly consistent in log space over all halo masses at the redshifts of interest ( $\sim \pm 0.15$  dex). As a result, varying  $\overline{SFR}(M_h, z)$  within the posterior can be approximated as a simple rescaling, and the effect of doing so is degenerate with that of  $\delta_{MF}$ . In the MCMC inference procedure (§2.3), we have absorbed the effect of this rescaling entirely into  $\delta_{MF}$ .

Figure 5 shows the effect of varying halo-to-halo scatter via  $\sigma_{SFR}$  and  $\sigma_{L_{CO}}$ . Increasing scatter increases the relative amplitude of the shot noise component of the power spectrum, compared to the clustering “bump” at low  $k$ . In the limit of high scatter,  $P(k) \sim \text{const.}$  or  $\Delta^2(k) \propto k^3$ . Notably, halo-to-halo scatter introduces sample variance in the signal *within the same volume*. However, this effect only appears to be significant for very high scatter:  $\sigma_{SFR} \sim 1.0$  dex or  $\sigma_{L_{CO}} \sim 1.0$ . These values are not expected to be realistic since they imply a 68% scatter of 2 dex.

We do not show the effect of varying the SFR-IR normalization  $\delta_{MF}$ , since it is simply a scaling factor: with other parameters fixed, a higher value of  $\delta_{MF}$  results in fainter CO brightness. For a given galaxy, a high value for  $\delta_{MF} (\gg 1)$  would mean a large amount of star formation per IR luminosity. This could indicate very low dust content or very short star formation timescales.

Figures 6 shows the effect of varying  $\alpha$  and  $\beta$  in the  $L_{\text{IR}} - L'_{\text{CO}}$  relation. Rewriting Eq. 2 as  $\log L'_{\text{CO}} = \alpha^{-1}(\log L_{\text{IR}} - \beta)$  makes the effect of each parameter more apparent. Independently increasing  $\alpha$  or  $\beta$  decreases  $L'_{\text{CO}}$  for a given  $L_{\text{IR}}$ . However, increasing  $\alpha$  also weights halos with low  $L_{\text{IR}}$  ( $\sim$  low SFR) more *relative* to halos with high  $L_{\text{IR}}$  ( $\sim$  high SFR). Depending on which halos are occupied by high-SFR galaxies, this can affect the shape of the power spectrum, particularly the clustering signal. This is a notable difference that can arise in predictions of power spectra if SFR is not modeled as simply linear in  $M_h$  (in our model, the average SFR turns downward just above  $M_h \sim 10^{12} M_\odot$ , a consequence of the now well-established downturn in the ratio of stellar mass to halo mass at a given halo mass).

### 3.3. MCMC Inference

Figure 7 shows the posterior distributions for all model parameters, as inferred from the “fiducial” and “non-detection” power spectra. In general, the full experiment places stronger constraints on the model parameters than the pathfinder.

Figure 8 shows constraints on the  $L_{\text{IR}} - L'_{\text{CO}}$  relation and the CO luminosity function, from both the pathfinder and full experiment observing our fiducial signal.

In the  $L_{\text{IR}} - L'_{\text{CO}}$  plots, we have overplotted the four empirical fits mentioned in §2.2.4 (Daddi et al. 2010; Carilli & Walter 2013; Greve et al. 2014; Dessauges-Zavadsky et al. 2015). The inferred relation is consistent with the fiducial line (Carilli & Walter 2013) within  $1\sigma$ , and consistent with all four lines within  $2\sigma$ . The notable point here is that while the four empirical fits were extrapolated from a limited, bright sample of galaxies (hence their convergence around  $\sim 10^9 - 10^{10}$  K km s<sup>-1</sup> pc<sup>2</sup>), the relation inferred from the intensity map was *constrained by probing faint populations directly*, albeit in integrated emission.

In the plots of CO luminosity function (LF), we recover the “true” CO luminosity function (underlying the fiducial intensity map) to within  $1\sigma$  over nearly all luminosities. Here, the full experiment produces tighter constraints than the pathfinder experiment, as expected. Note that the “true” LF

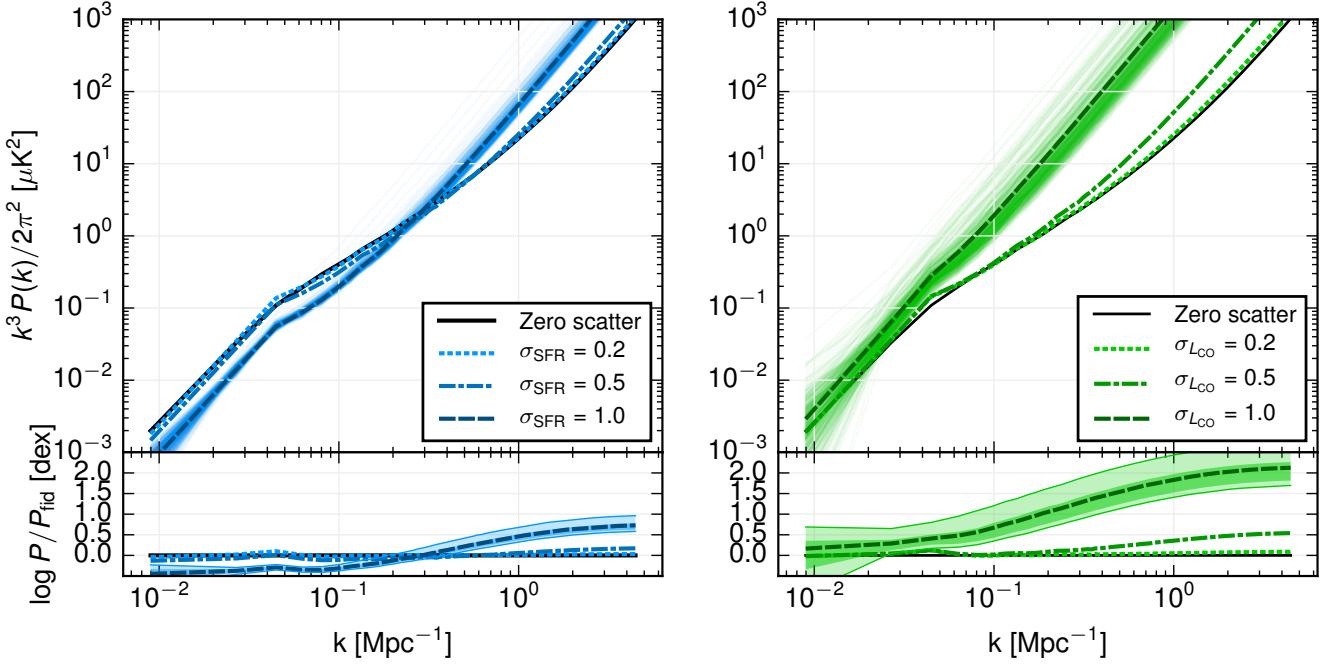


FIG. 5.— Effect on the CO power spectrum of halo-to-halo scatter as parameterized by: SFR given  $M_h$  ( $\sigma_{\text{SFR}}$ , left) or  $L_{\text{CO}}$  given SFR ( $\sigma_{L_{\text{CO}}}$ , right). At  $\sigma_{\text{SFR}}, \sigma_{L_{\text{CO}}} \approx 1$  dex and higher, two things are evident: (1) the power spectrum begins to look like a pure shot noise spectrum, since any clustering signature is increasingly buried by the large halo-to-halo scatter, and (2) the scatter introduces significant variance into the power spectrum. This variance is demonstrated for the  $\sigma = 1.0$  case by plotting 100 individual power spectra in the top plot, as well as their shaded 95% interval in the bottom plot. However, this scenario is probably extreme and unrealistic, because it implies a  $\pm 1\sigma$  scatter over two orders of magnitude. In this plot, the labeled lines are the mean of 100 power spectra from the exact same halos (rather than multiple realizations of the survey volume) to isolate the variance introduced by halo-to-halo scatter alone.

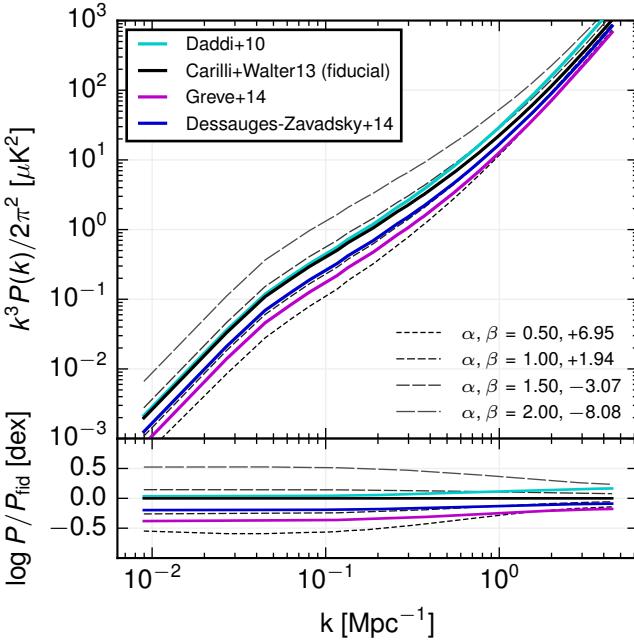


FIG. 6.— Power spectra due to varying  $L_{\text{IR}}-L_{\text{CO}}$  relations. Solid colored lines show literature values of  $\alpha, \beta$  (Daddi et al. 2010; Carilli & Walter 2013; Greve et al. 2014; Dessauges-Zavadsky et al. 2015). Gray dashed lines show the power spectra from  $\alpha = 0.5, 1.0, 1.5, 2.0$  (dashed lines) and corresponding  $\beta$  assuming  $\alpha$  and  $\beta$  covary as in Eq. 3. As  $\alpha$  increases, smaller and fainter (in IR) galaxies contribute a greater proportion of the total CO luminosity, and the relative strength of the low- $k$  clustering signal increases.

lies outside of the  $1\sigma$  interval at the bright end. While this is not a very strong deviation, we expect that intensity mapping will not be especially sensitive to the bright end of the LF if

somewhat fainter populations dominate the overall signal, as they do here.

We also show  $2\sigma$  posterior constraints on the CO luminosity function in the event that no clear signal is detected (Fig. 8, bottom right). While these constraints visually resemble upper limits, they are actually allowed regions and are therefore sensitive to our choice of priors. We include further discussion of these priors in §4.3.

To compare the LF constraints with a hypothetical blind survey for CO in individual galaxies, we have marked the faintest CO luminosity that could be detected by the VLA in 1 and 100 hr of on-source time. For the calculation, we used the VLA exposure calculator<sup>8</sup>, assuming an observing frequency of 32 GHz ( $z \approx 2.6$ ), 100 km s $^{-1}$  channel, and 5 $\sigma$  detection threshold. 100 hours of on-source time can probe below the “knee” of the luminosity function predicted by this model. However, the relatively small field of view ( $\sim 1.3'$  primary beam at 32 GHz) means that sample variance in such a blind survey will be significant.

### 3.3.1. Cross-check with Mock Data from Obreschkow et al. (2009)

The model used to generate the mock signal is the same model used in the MCMC procedure. It is reassuring, but perhaps unsurprising, that we recover the “true” CO luminosity function when fitting this model. What if we use a population of CO galaxies that was generated from an independent model applied to a different simulation?

Various studies have provided physically motivated modeling of  $L_{\text{CO}}$  at high redshift, whether through direct hydrodynamic simulation or semi-empirical modeling (e.g. Popping et al. 2014; Lagos et al. 2015). For this exercise, meant to pro-

<sup>8</sup> <https://obs.vla.nrao.edu/ect>

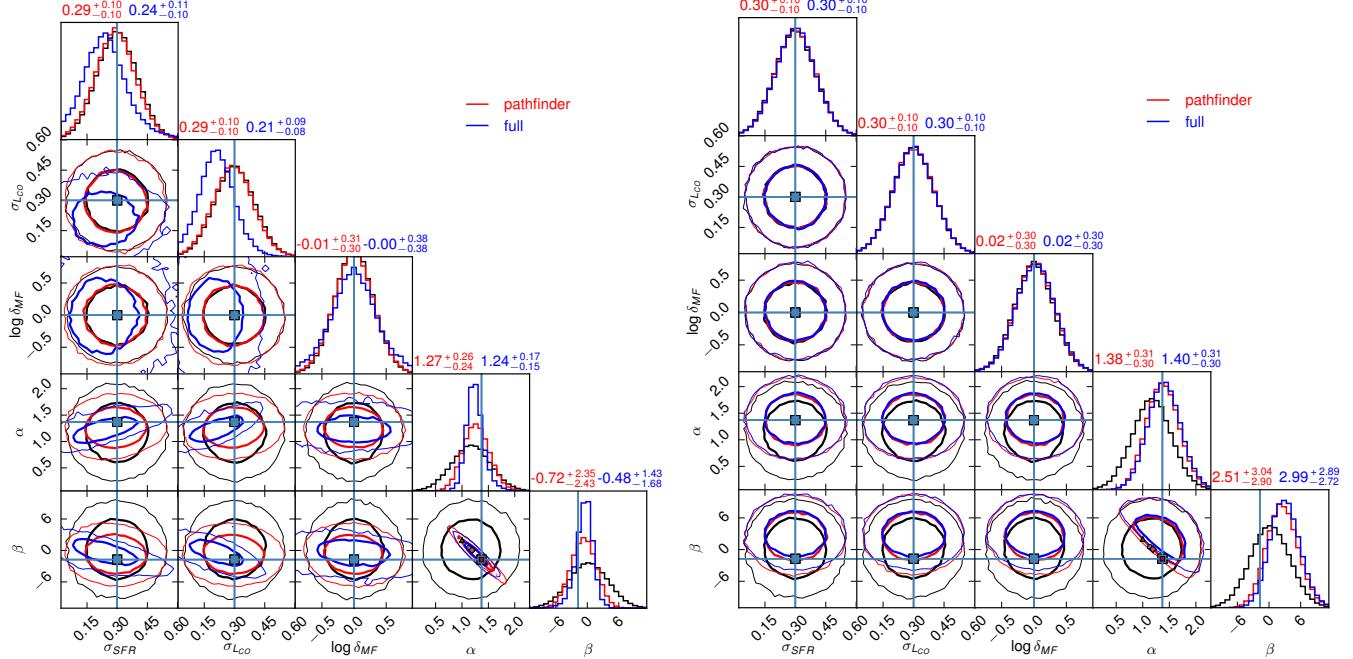


FIG. 7.— MCMC posterior distributions on the parameter space of our model. Black contours show prior distributions. Red contours show constraints from the pathfinder experiment, while blue contours show constraints from the full experiment. Crosshairs indicate the values of our fiducial model. Marked points in the  $\alpha$ - $\beta$  plot (bottom row, second from right) show values obtained by previous studies (see §2.2.4). **Left:** Posterior distributions inferred from our fiducial signal, as observed by both the pathfinder (red) and full (blue) experiments. **Right:** Posterior distributions inferred from a non-detection by the same two experiments.

vide a sanity check rather than a prediction, we assume the underlying galaxy  $L_{\text{CO}}$ 's are given by the results of Obreschkow et al. (2009a,b), who use a semi-analytic model to predict galaxy CO luminosities from dark matter halos. From their publicly downloadable lightcone data<sup>9</sup>, we generate CO intensity maps and rerun the same analysis, for both the pathfinder and full experiments. See Appendix E for more details.

Figure 9 compares the actual CO luminosity function in the volume with the inferred luminosity function from intensity mapping. The weaker signal is consistent with non-detection for the pathfinder. However, it is detected by the full experiment, and within the 95% credible intervals, the inferred LF is consistent with the "true" LF, while our original model's LF is largely excluded. While this is not an exhaustive check, it is an indication that our model appears flexible enough to accommodate a reasonable range of predicted LFs.

#### 4. DISCUSSION

##### 4.1. Relation to Previous Work

Our work here naturally extends the various studies mentioned in §2.1. In modeling, our main improvements are (1) an empirically constrained connection between halos and mean SFRs and (2) an  $L_{\text{IR}}-L_{\text{CO}}$  connection that is fit to a larger set of observations. These should model more accurately the connection between underlying large-scale structure and the predicted CO intensity map. Qualitatively, the shape of the overall  $L_{\text{CO}}(M_h)$  relation is somewhat similar to that of Gong et al. (2011), though their focus was on  $z \gtrsim 6$ .

Figure 10 shows the power spectrum of our fiducial model compared to other models at similar redshifts (see Fig. 2). Note that the power spectra of those models were calculated with a halo mass cutoff of  $M_{\text{CO}, \min} = 10^9 M_\odot$ , lower than our fiducial  $10^{10} M_\odot$ , to better compare with their original pre-

dictions. However, this does not affect the main qualitative point of the plot: the nonlinear form of our  $L_{\text{CO}}(M_h)$  affects the shape of the power spectrum — in this case, increasing the relative power in clustering (low- $k$ ) modes.

The novel aspect of this work is an attempt to use CO intensity mapping — or rather, simulated observations thereof — to make inferences about the underlying galaxy population. This is particularly significant because, despite the impressive capabilities of ALMA and VLA, it will still be difficult for them to probe complete samples of the faint end of the CO luminosity function. Considering that CO is the most accessible tracer of molecular gas, relying solely on galaxy surveys may limit our ability to trace the gas-SFR connection at high redshifts.

In our approach, we have used full  $N$ -body simulation, which is more computationally expensive than analytically approximating the power spectrum, even if existing halo catalogs for a simulation are already available. Indeed, with a simple  $L_{\text{CO}} \propto M$  model, the power spectrum can be quickly and analytically calculated from the halo model formalism (Cooray & Sheth 2002). This is how the  $L_{\text{CO}} \propto M$  power spectra (i.e. previous models) of Figure 10 were calculated. However, directly simulating these halos is a method that more naturally accommodates information about galaxy environment and merger histories. For example, merger-driven starbursts could be identified and modeled separately from the "normal" star-forming population. Modeling of that nature is left for future studies. Galaxy formation involves many complex processes with diverse observational signatures, and this approach provides a framework for all of those facets to be studied more directly.

##### 4.2. Modeling Uncertainties

One goal of intensity mapping is to probe high-redshift galaxies too faint to be individually observed in large num-

<sup>9</sup> [http://s-cubed.physics.ox.ac.uk/s3\\_sax](http://s-cubed.physics.ox.ac.uk/s3_sax)

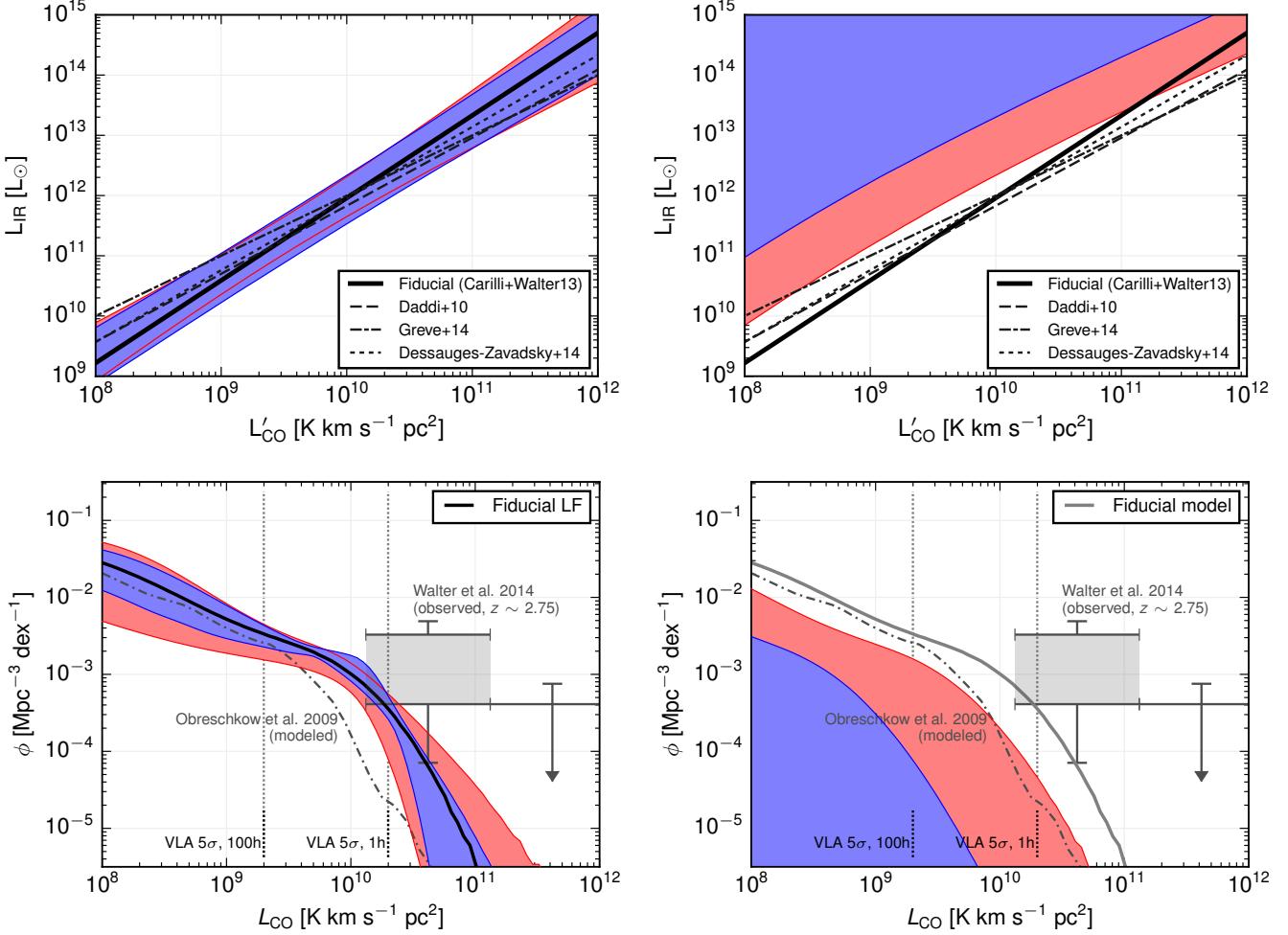


FIG. 8.— The properties of galaxy populations, as inferred by the pathfinder (red) and full (blue) experiments. **Top left:** Inferred  $1\sigma$  constraints on the mean  $L_{\text{IR}} - L'_{\text{CO}}$  relation. Note that while the pathfinder may appear to yield comparable constraints to the full experiment, this is an artifact of our somewhat simple priors and likelihood (see §3.3). The parameters  $\alpha$  and  $\beta$  are individually more constrained by the full experiment (see Fig 7). For comparison, we also plot derived fits from [Daddi et al. \(2010\)](#), [Greve et al. \(2014\)](#), and [Dessauges-Zavadsky et al. \(2015\)](#). **Top right:** 95% posterior constraints for the  $L_{\text{IR}} - L'_{\text{CO}}$  relation from a non-detection. **Bottom left:** Inferred CO luminosity functions. Colored intervals show the median and 95% credible interval for inferred CO luminosity functions. For comparison, we show observational constraints at  $z \sim 2.75$  from a blind scan of the Hubble Deep Field ([Walter et al. 2014](#)), as well as the model predictions of [Obreschkow et al. \(2009a\)](#) at similar redshifts. The vertical dashed lines marks the VLA  $5\sigma$  detection threshold assuming 1 and 100 hr of on-source time and a 100 \$\text{km s}^{-1}\$ channel width at 32 GHz. **Bottom right:** 95% posterior constraints on the CO luminosity function from a non-detection.

bers. By nature, then, these are currently poorly studied populations, and any attempt to empirically model their  $L_{\text{CO}}$ 's must rely on data from (1) galaxies thought to be low-redshift analogs, e.g. local dwarfs and spirals; (2) observed galaxies at high redshift, which may be atypically luminous or star-forming; or (3) galaxies that might be “typical” at high redshift, but in small and incomplete samples.

Predictions about such galaxies are subject to significant uncertainties about the formation and evolution of high-redshift galaxies. In particular, the nature of the ISM at high redshift is not fully understood. In the very early universe, one might expect significantly lower metallicity and dust shielding in the ISM, resulting in lower CO abundance. However, in active star-forming regions, young massive stars should quickly pollute their surroundings with metals. These various effects have not been well quantified in typical high-redshift galaxies and are just beginning to be probed by individual galaxy observations with, e.g., ALMA and VLA.

#### 4.2.1. Using $L_{\text{IR}}$ to trace SFR and $L_{\text{CO}}$

In our model,  $L_{\text{IR}}$  correlates with SFR (Eq. 1). This is because we need an empirical tracer of star formation to link halo SFR with CO luminosity. Using  $L_{\text{IR}}$  assumes that star formation occurs in the presence of dust, and that  $L_{\text{IR}}$  is thermally radiated from dust heated by massive, young stars.

The normalization  $\delta_{\text{MF}}$ , as well, is sensitive to star formation history and the initial mass function, both of which may be different at high redshift. It should also be noted that a tracer like  $L_{\text{IR}}$  may overestimate the instantaneous SFR if star formation occurs in short “bursty” episodes ( $\ll 100$  Myr, e.g. [Domínguez et al. 2015](#)) or in recently quenched galaxies ([Hayward et al. 2014](#)).

Additionally, although we have taken  $L_{\text{IR}}$  to be the total luminosity in the  $8 - 1000 \mu\text{m}$  band, note that this is not always the directly observed quantity.  $L_{\text{IR}}$  may be converted from a measurement of flux in a narrower IR band. Any systematic errors introduced by this fact will probably not be dominant, but it is something to remember going forward.

#### 4.2.2. CO Line Luminosities

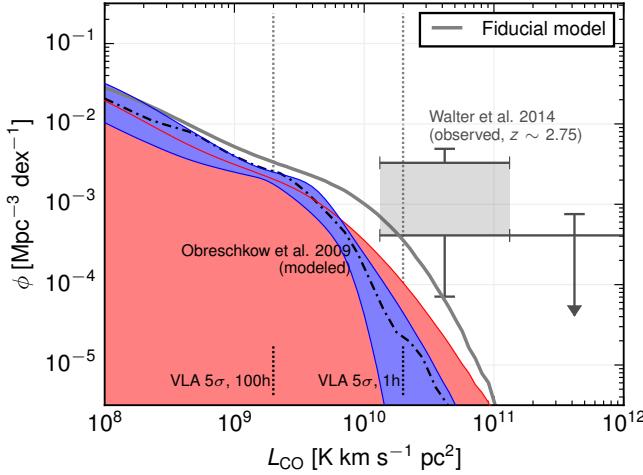


FIG. 9.— CO luminosity function from the data of Obreschkow et al. (2009b), as inferred by the pathfinder and full experiments. Our fiducial model has been plotted for comparison, but note that the inferred LF matches the Obreschkow et al. (2009b) curve more closely now while excluding our fiducial model (as should be expected). For more details, see Fig. 8.

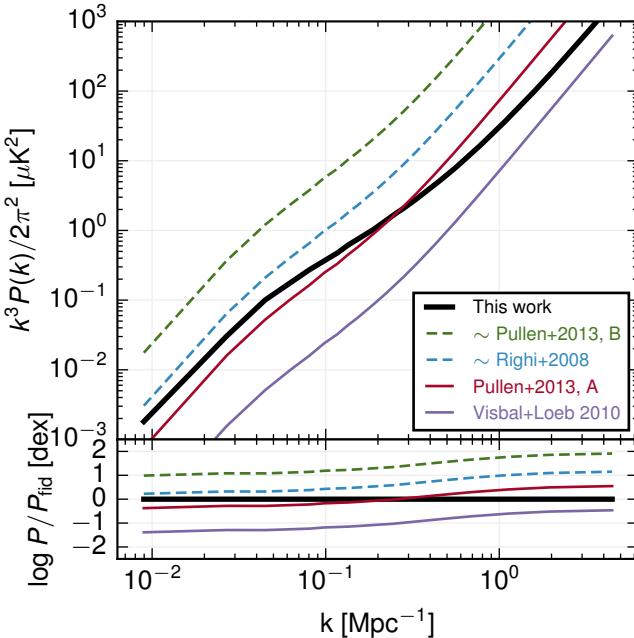


FIG. 10.— Comparison of the fiducial power spectrum in this work with that of previous models (Righi et al. 2008; Visbal & Loeb 2010; Pullen et al. 2013, see Fig. 2), which either modeled  $L_{\text{CO}} \propto M_h$  or have been approximated as such here. Our model does not have a linear  $L_{\text{CO}}(M_h)$ , and this qualitatively affects the shape of the resulting power spectrum. Nevertheless, our fiducial prediction appears to be consistent with the range of previous predictions.

Existing high-redshift data for CO(1-0) luminosity are also not free of uncertainty. Some values for the CO(1-0) luminosity were not directly measured, but instead inferred from higher-order transitions that were accessible in an available observing band (e.g. Genzel et al. 2010; Tacconi et al. 2013). Inferring CO(1-0) luminosities from such observations relies on assumptions about spectral line energy distribution (SLED), i.e. the relative luminosities of each CO line. In the optically thick, high-temperature limit, all lines have the same brightness temperature, or equivalently luminosities of  $L_{\text{CO}(J \rightarrow J-1)} \propto J^2 L_{\text{CO}(1-0)}$ , but this generally overestimates

high- $J$  luminosities relative to  $L_{\text{CO}(1-0)}$ .

While some recent work has focused on characterizing the SLED across a sample of galaxies (e.g. Greve et al. 2014; Narayanan & Krumholz 2014), the scatter is large and predicting the SLED in an individual galaxy is not straightforward. On the other hand, it may be possible to constrain the average SLED by cross-correlating two or more intensity maps that target the same cosmological volume through different CO transitions.

#### 4.3. Inferred Constraints on Galaxy Populations

The  $L_{\text{IR}}-L'_{\text{CO}}$  relation (Eq. 2) has been fit to observations in previous literature (Daddi et al. 2010; Dessauges-Zavadsky et al. 2015; Carilli & Walter 2013; Greve et al. 2014). In Figure 8, we have plotted this relation, both as derived in the literature and as inferred from a hypothetical CO signal. To the extent that the  $L_{\text{IR}}-L'_{\text{CO}}$  relation traces the star formation–molecular gas relation, it is possible for intensity mapping to characterize the latter in unobserved galaxy populations.

Our results also constrain the CO luminosity function. Currently, this is not well characterized at the redshifts in this study because of the long integration times required to detect CO in a single galaxy. We have overplotted the results of Walter et al. (2014), which estimated the CO luminosity function from a blind scan for CO lines in the Hubble Deep Field.

The noteworthy point here is that these constraints, especially for the faintest galaxies, arise from directly imaging their emission, albeit in aggregate. By contrast, because galaxy surveys preferentially probe the brightest galaxies, inferences about the faintest galaxies need to be made by, e.g., extrapolation. With intensity mapping, part of the information we receive comes directly from the (integrated) faint end of the luminosity function.

In future analyses, one can certainly consider tighter priors than we have chosen here. In the event of a *detection* of an intensity mapping signal, we expect narrower priors to yield narrower constraints on the CO luminosity function. In the event of a *non-detection*, we would actually expect narrower priors to yield less stringent constraints, i.e. a higher upper bound on the allowed luminosity function. This is because our broad priors on  $\alpha$  and  $\beta$  (of the  $L_{\text{IR}}-L_{\text{CO}}$  relation) actually allow a significant region of parameter space (high  $\alpha$  and  $\beta$ ) that would be in significant tension with existing  $L_{\text{IR}}-L_{\text{CO}}$  data, even though it yields undetectable CO power spectra, consistent with the “observed” intensity map.<sup>10</sup> As a result, given a mock non-detection, a region of parameter space producing unrealistically faint populations is still allowed in the posterior region. Tighter priors would exclude that unrealistic parameter space, raising the upper bound on the allowed luminosity function. All the same, the plots in Fig. 8 do demonstrate that a non-detection of an intensity mapping signal can constrain the faint end of the luminosity function. Alternatively, a more complete likelihood calculation—for example, simultaneously fitting the  $L_{\text{IR}}-L'_{\text{CO}}$  relation to galaxy survey and intensity mapping data—should also yield stronger and more robust constraints.

One parameter we have not allowed to vary is  $M_{\text{CO}, \min}$ , the minimum mass of CO-luminous halos, which we fix at  $10^{10} M_\odot$ . In principle, this could also affect the signal. However, as we note in Appendix A, lowering the mass cut-

<sup>10</sup> This also appears to be why the  $L_{\text{IR}}-L'_{\text{CO}}$  constraints in Fig. 8 do not seem to be improved by the full experiment, a somewhat counterintuitive result.

off yields diminishing returns in our fiducial model because  $L_{\text{CO}}(M)$  decreases toward low  $M$  faster than the halo number density increases. Any variation in the signal we might get from decreasing  $M_{\text{CO}, \min}$  below below  $10^{10} M_{\odot}$  is expected to be eclipsed by other uncertainties in the model. However, one can certainly imagine more realistic models than those using an abrupt halo mass cutoff. We leave investigation of these models to future studies.

A consideration in making meaningful inferences is that the spherically averaged power spectrum essentially carries two independent sources of information: (1) the amplitude of the clustering (low- $k$ ) signal, and (2) the amplitude of the shot noise (high- $k$ ) signal. This restricts the complexity of models we can expect to meaningfully constrain. However, there are other effects that are not modeled here but will be implemented as refinements in a future study. These include redshift space distortions and line broadening, which will distort the signal along the line of sight. This slightly alters the power spectrum in that direction, potentially providing additional information. Cross-correlation could also be useful in this regard, whether with intensity maps of other lines or galaxy surveys. The latter would include only the brightest galaxies but nevertheless trace the same cosmological structure.

It is worth noting that, although we have focused this analysis on CO, this same approach could just as well be applied to, say, [CII] intensity mapping. [CII] is also thought to trace star formation in high-redshift galaxies, but through a different (ionized) phase of the ISM.

#### 4.4. Astrophysical Contaminants and Complexities

In this paper, we have assumed for simplicity that the cosmological CO(1-0) map been perfectly isolated from the raw signal. Reality is certainly more complex.

Most of the astrophysical foregrounds and backgrounds are expected to be continuum emission. These include the CMB itself, synchrotron radiation within the Milky Way, and (redshifted) thermal dust emission from all galaxies along the line of sight. These components are expected to be relatively smooth in frequency space, so if the instrument has enough spectral resolution, the continua can plausibly be subtracted from the signal, while the remaining fluctuations should contain the cosmological CO signal. Note that this continuum subtraction effectively subtracts the lowest- $k$  modes ( $k = 0$  is the mean intensity), and so information in those modes will be reduced or removed. If those particular modes are important for the particular analysis at hand, the survey would need to be expanded to a larger volume to recover them.

Line broadening and redshift space distortions, while not contaminants, will alter the power spectrum along the line of sight. This and previous studies have assumed a delta-function line profile, with zero width. However, if CO is a reasonable dynamical tracer within galaxies, then line broadening is an important consideration. We would expect broadening to reduce small-scale (high- $k$ ) power along the line of sight, limiting the narrowest useful frequency channel. For reference, a rest-frame width of  $100 \text{ km s}^{-1}$  for CO(1-0) from  $z \approx 2.5$  is an observed frequency width of  $11 \text{ MHz}$  (roughly the frequency resolution of the “full” experiment in this paper; compare  $8.25 \text{ MHz} \sim 1 \text{ Mpc}$  from §2.4). Additionally, galaxy peculiar velocities will cause redshift space distortions, altering the apparent clustering of CO emission. Properly accounting for these effects requires the cylindrical power spectrum  $P(k_{\perp}, k_{\parallel})$  instead of the spherical  $P(k)$  in this study.

The signal may be partially contaminated by interloping non-CO spectral lines that have been redshifted into the observed frequency band. While we expect low- $J$  CO lines to be the brightest lines in their spectral neighborhood, other lines may cumulatively add spurious fluctuations on top of the CO signal. In particular, HCN ( $\nu_{\text{rest}} = 88.63 \text{ GHz}$  from  $z \sim 1.6 - 2.0$ ) may contribute non-negligible shot noise to the measured power spectrum. In a 2D  $C_{\ell}$  analysis, Breysse et al. (2015) found that using pixel masking to mitigate HCN foregrounds resulted in a loss of shot noise information. As inferences about galaxy populations are sensitive to both clustering and shot noise components, this could complicate the astrophysical interpretation of an intensity mapping measurement. We note that, while data for the intensities of contaminant lines are currently sparse, in ALMA observations of lensed high-redshift star-forming galaxies, HCN is dimmer than CO by at least an order of magnitude (Spilker et al. 2014).

Our observed band should also catch CO(2-1) from  $z \approx 5.8 - 6.7$ . We do not expect it to dominate the CO(1-0) signal. However, removing the CO(1-0) signal could give us a residual CO(2-1) signal, if this removal can be done robustly enough, e.g. by cross-correlating with galaxy surveys. This would allow one band to simultaneously study two redshift ranges.

Note that we do not expect the bright  $158 \mu\text{m}$  [CII] line to be a contaminant. For the observing band considered in this study, an interloping [CII] line would have to originate from  $z \gtrsim 55$ , when the universe is  $< 50 \text{ Myr old}$ . It is unlikely that [CII] will be found in significant abundance at these redshifts (though it would be a remarkable discovery).

Finally, one more concern is that since the CMB would have a higher temperature at high redshift, then as an observing background it may reduce the apparent luminosity of the CO lines (Obreschkow et al. 2009a; da Cunha et al. 2013). The severity of this effect depends on the ISM properties of high-redshift galaxies: if the CO-luminous regions of the ISM are significantly warmer at high redshifts, the CMB is less of a concern, and there is some evidence that this may be true (Harris et al. 2012; Daddi et al. 2015). However, suppose the CO(1-0) luminosity is globally reduced by a factor 0.4, which is similar to or more pessimistic than the predictions in da Cunha et al. (2013) at  $z \sim 2 - 3$ .  $P(k)$  would be reduced by a factor 0.16, and recalculating the fiducial detection significance  $\text{SNR}_{\text{tot}}$  (§3.1, Eq. 7) yields  $\sim 1.67\sigma$  and  $49.6\sigma$  for the pathfinder and full experiments, respectively. While these numbers are certainly smaller, they still suggest a signal detectable by the full experiment. Nevertheless, the effect of the CMB as an observing background remains a valid concern due to unresolved uncertainties about the high-redshift ISM.

#### 4.5. Prospects for Cross-correlation

Because intensity mapping smoothes over emission from many galaxies, cross-correlation of the signal with galaxy surveys offers a promising method of isolating the signal from galaxies alone (e.g. Pullen et al. 2013) at a specific redshift. The underlying idea is not that CO would necessarily originate only from observed galaxies, but rather that individual galaxies and CO would trace the same large-scale structure. As a specific example, the survey area of the pathfinder experiment is well matched to certain galaxy surveys, including the  $2 \text{ deg}^2$  COSMOS field (Scoville et al. 2007), which contains extensive multiwavelength data. Cross-correlating the CO map with existing galaxies can help validate the signal and may also constrain the properties of undetected galaxies.

Within the same field, multiple intensity maps of different lines could be quite complementary. CO is particularly well suited for cross-correlation with itself since it emits in a known ladder of emission lines, so, for example, CO(1-0) emission observed at  $\nu_{\text{obs}}$  should correlate with CO(2-1) emission at  $2\nu_{\text{obs}}$ . In addition, cross-correlation between CO and [CII] signals from the same redshift, as well as between CO and H $\alpha$  maps, can provide crucial information on the nature of various phases of the ISM in galaxies.

Simultaneous intensity maps of H $\alpha$  and CO emission, in particular, would be especially useful in probing the reionization epoch (e.g. Lidz et al. 2011). The former would trace the process of reionization, while the latter might trace star-forming galaxies. If reionization is in fact driven by star-forming galaxies, cross-correlating the two maps could place interesting constraints on the morphology of the reionized intergalactic medium (via H $\alpha$ ), the nature of the reionizing sources (via CO), and the relationship between the two.

## 5. SUMMARY

We have performed a preliminary analysis of the ability of CO intensity mapping to probe high-redshift galaxy populations. The following list summarizes our main results:

1. Based on our fiducial assumptions, we find that **the CO(1-0) signal from  $z \sim 2.4\text{--}2.8$  should be detectable by a realistic intensity mapping experiment** (§3.1). Our model is consistent with a range of previous predictions, but all models carry significant uncertainties, which are not definitively resolved by current observations.
2. **Details of the  $L_{\text{CO}}\text{--}M_{\text{h}}$  relation produce measurable differences in shape of the power spectrum** (§3.2). This encodes information about the galaxy population in the relative strengths of clustering (low- $k$ ) and shot noise (high- $k$ ) modes. This implies that sensitivity to fluctuations over a large dynamic range of scales provides better constraints on the underlying galaxy population.

## REFERENCES

Anderson, L., Aubourg, E., Bailey, S., et al. 2012, MNRAS, 427, 3435 [§ 1]  
 Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33 [§ 5]  
 Bandura, K., Addison, G. E., Amiri, M., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9145, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 22 [§ 1]  
 Barger, A. J., Cowie, L. L., & Richards, E. A. 2000, AJ, 119, 2092 [§ 2.2.3]  
 Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013a, ApJ, 770, 57 [§ 2.2.2, 2.2.3, 1, 4, 3.2, 11, 12, A.1, A.2]  
 —. 2013b, ApJ, 762, L31 [§ 2.2.2]  
 Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013c, ApJ, 762, 109 [§ 2.2.1]  
 Béthermin, M., Daddi, E., Magdis, G., et al. 2015, A&A, 573, A113 [§ 2.2.3]  
 Blake, C., Kazin, E. A., Beutler, F., et al. 2011, MNRAS, 418, 1707 [§ 1]  
 Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, ARA&A, 51, 207 [§ 1]  
 Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2014, MNRAS, 443, 3506 [§ 1, 2, 2.1, 2.4.1, 3.1, 3.1, D]  
 —. 2015, MNRAS, 452, 3408 [§ 4.4]  
 Bunker, A. J., Wilkins, S., Ellis, R. S., et al. 2010, MNRAS, 409, 855 [§ 3]  
 Carilli, C. L. 2011, ApJ, 730, L30 [§ 1, a, 2.1, 2.2.3, 2.4.1, 3]

Carilli, C. L., & Walter, F. 2013, ARA&A, 51, 105 [§ 2.2.3, 2.2.4, 3.3, 6, 4.3]  
 Chabrier, G. 2003, PASP, 115, 763 [§ 2.2.3]  
 Chang, T.-C., Pen, U.-L., Peterson, J. B., & McDonald, P. 2008, Physical Review Letters, 100, 091303 [§ 1]  
 Cooray, A., & Sheth, R. 2002, Phys. Rep., 372, 1 [§ 3.1, 4.1]  
 Crocce, M., Pueblas, S., & Scoccimarro, R. 2006, MNRAS, 373, 369 [§ 2.2.1]  
 da Cunha, E., Groves, B., Walter, F., et al. 2013, ApJ, 766, 13 [§ 4.4]  
 Daddi, E., Elbaz, D., Walter, F., et al. 2010, ApJ, 714, L118 [§ 2.2.4, 3.3, 6, 8, 4.3, 3]  
 Daddi, E., Dannerbauer, H., Liu, D., et al. 2015, A&A, 577, A46 [§ 4.4]  
 Dessauges-Zavadsky, M., Zamojski, M., Schaerer, D., et al. 2015, A&A, 577, A50 [§ 2.2.4, 5, 3.3, 6, 8, 4.3]  
 Domínguez, A., Siana, B., Brooks, A. M., et al. 2015, MNRAS, 451, 839 [§ 4.2.1]  
 Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306 [§ 2.3, B.3]  
 Foreman-Mackey, D., Price-Whelan, A., Ryan, G., et al. 2014, doi:10.5281/zenodo.11020 [§ 5]  
 Genzel, R., Tacconi, L. J., Gracia-Carpio, J., et al. 2010, MNRAS, 407, 2091 [§ 4.2.2, 12, A.2]  
 Gong, Y., Cooray, A., Silva, M., et al. 2012, ApJ, 745, 49 [§ 1, C]  
 Gong, Y., Cooray, A., Silva, M. B., Santos, M. G., & Lubin, P. 2011, ApJ, 728, L46 [§ 1, a, 2.1, 4.1, 3]

<sup>11</sup> <http://www.astropy.org>

3. **We demonstrate the extent to which an intensity mapping observation can constrain the properties of galaxy populations** (§3.3), namely (1) the  $L_{\text{IR}}\text{--}L_{\text{CO}}$  relation and (2) the CO luminosity function. This has significance for understanding molecular gas and its connection to star formation in high-redshift galaxies, to the extent that  $L_{\text{IR}}$  traces SFR and  $L_{\text{CO}}$  traces molecular gas. At high redshifts, those connections will require further study.

We reiterate that intensity mapping directly probes the faint end of the luminosity function while similar inferences from galaxy surveys rely on extrapolation from observations of bright galaxies. This suggests that intensity mapping and galaxy surveys could be complementary avenues for understanding galaxy populations at high redshifts.

In summary, we have shown that intensity mapping is a promising method of observing CO in high-redshift galaxies. If CO traces molecular gas in these galaxies, then CO intensity mapping is a potentially informative probe of star formation and molecular gas in the high-redshift universe, particularly in galaxies that will be hard to observe individually for the foreseeable future.

T.Y.L. and R.H.W. received support from the U.S. Department of Energy under contract number DE-AC02-76SF00515 as well as from KIPAC Enterprise funds. This work made use of computational resources at the SLAC National Accelerator Laboratory. We thank Matthew Becker for providing access to the cosmological simulation ( $c400\text{--}2048$ ) used in this work. We have also used or modified the following public Python packages: Astropy<sup>11</sup> (Astropy Collaboration et al. 2013); triangle\_plot (Foreman-Mackey et al. 2014, for Fig. 7); and hmf (Murray et al. 2013, for Figs. 10 and 11).

We are grateful to Marcelo Alvarez, Peter Behroozi, Kieran Cleary, Olivier Doré, Andy Harris, Yashar Hezaveh, Ryan Keisler, Daniel Marrone, Gergö Popping, Anthony Pullen, Tony Readhead, Marco Viero, and members of the COMAP collaboration for various helpful discussions and/or comments on a previous draft. Finally, we would like to thank an anonymous referee for their constructive comments which improved this paper.

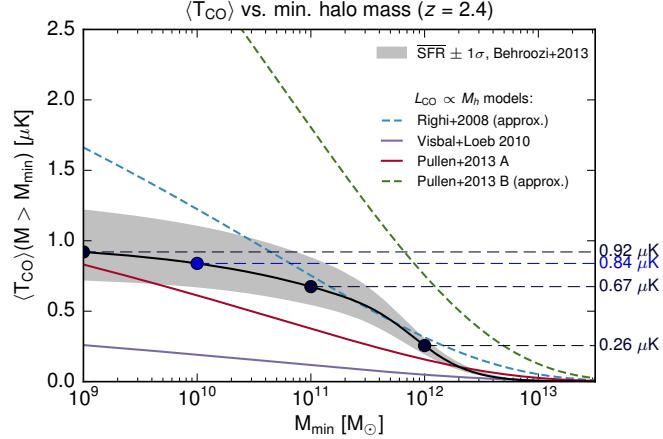


FIG. 11.— Mean CO brightness temperature as a function of minimum halo mass, at  $z = 2.4$ . As  $M_{\text{min}}$  increases,  $\langle T_{\text{CO}} \rangle$  naturally decreases, since fewer halos contribute to total CO luminosity. The shaded gray range indicates the values spanned by the  $\langle T_{\text{CO}} \rangle$  spanned by the  $\pm 1\sigma$  posteriors on  $\text{SFR}(M, z)$  from Behroozi et al. (2013a). The values of  $\langle T_{\text{CO}} \rangle$  at  $M_{\text{min}} = 10^9, 10^{10}, 10^{11}$ , and  $10^{12} M_{\odot}$  have been labeled on the right. Note that  $\langle T_{\text{CO}} \rangle$  was analytically calculated from the mass function of Sheth et al. (2001).

Goodman, J., & Weare, J. 2010, CAMCoS, 5, 65 [§ 2.3, B.3]  
 Greve, T. R., Leonidaki, I., Xilouris, E. M., et al. 2014, ApJ, 794, 142 [§ 2.2.4, 3.3, 6, 8, 4.2.2, 4.3]  
 Harris, A. I., Baker, A. J., Frayer, D. T., et al. 2012, ApJ, 752, 152 [§ 4.4]  
 Hayward, C. C., Lanz, L., Ashby, M. L. N., et al. 2014, MNRAS, 445, 1598 [§ 4.2.1]  
 Kennicutt, Jr., R. C. 1998, ApJ, 498, 541 [§ 2.2.3, 2.2.3, 3]  
 Lacey, C., & Cole, S. 1993, MNRAS, 262, 627 [§ 3]  
 Lagos, C. d. P., Crain, R. A., Schaye, J., et al. 2015, MNRAS, 452, 3815 [§ 3.3.1]  
 Lidz, A., Furlanetto, S. R., Oh, S. P., et al. 2011, ApJ, 741, 70 [§ 1, a, 2.1, 2.2.3, 4.5, C, C.3, 3]  
 Madau, P., Haardt, F., & Rees, M. J. 1999, ApJ, 514, 648 [§ 3]  
 Magnelli, B., Lutz, D., Santini, P., et al. 2012, A&A, 539, A155 [§ 2.2.3]  
 Morales, M. F., & Wyithe, J. S. B. 2010, ARA&A, 48, 127 [§ 1]  
 Murray, S. G., Power, C., & Robotham, A. S. G. 2013, Astronomy and Computing, 3, 23 [§ 5]  
 Narayanan, D., & Krumholz, M. R. 2014, MNRAS, 442, 1411 [§ 4.2.2]  
 Obreschkow, D., Heywood, I., Klöckner, H.-R., & Rawlings, S. 2009a, ApJ, 702, 1321 [§ 3.3.1, 8, 4.4, E, 3]  
 Obreschkow, D., Klöckner, H.-R., Heywood, I., Levrier, F., & Rawlings, S. 2009b, ApJ, 703, 1890 [§ 3, 2.3, 3.3.1, 9, E]  
 Omont, A., Cox, P., Bertoldi, F., et al. 2001, A&A, 374, 371 [§ 2.2.3]  
 Popping, G., Somerville, R. S., & Trager, S. C. 2014, MNRAS, 442, 2398 [§ 3.3.1]  
 Pullen, A. R., Chang, T.-C., Doré, O., & Lidz, A. 2013, ApJ, 768, 15 [§ 1, 2.1, 2.2.3, 2.2.4, 2.4.1, 3.1, 10, 4.5, 3]  
 Pullen, A. R., Doré, O., & Bock, J. 2014, ApJ, 786, 111 [§ 1]  
 Riechers, D. A., Carilli, C. L., Capak, P. L., et al. 2014, ApJ, 796, 84 [§ 1]

Righi, M., Hernández-Monteagudo, C., & Sunyaev, R. A. 2008, A&A, 489, 489 [§ 2, 2.1, 10, 3]  
 Rowan-Robinson, M. 2000, MNRAS, 316, 885 [§ 2.2.3]  
 Salmon, B., Papovich, C., Finkelstein, S. L., et al. 2015, ApJ, 799, 183 [§ 2.2.2]  
 Scoville, N., & Young, J. S. 1983, ApJ, 265, 148 [§ 2.2.3]  
 Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1 [§ 4.5]  
 Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1 [§ 11, A.1]  
 Silva, M., Santos, M. G., Cooray, A., & Gong, Y. 2015, ApJ, 806, 209 [§ 1]  
 Smit, R., Bouwens, R. J., Franx, M., et al. 2012, ApJ, 756, 14 [§ 3]  
 Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15 [§ 2.2.2]  
 Spilker, J. S., Marrone, D. P., Aguirre, J. E., et al. 2014, ApJ, 785, 149 [§ 4.4]  
 Springel, V. 2005, MNRAS, 364, 1105 [§ 2.2.1]  
 Springel, V., Yoshida, N., & White, S. D. M. 2001, New A, 6, 79 [§ 2.2.1]  
 Tacconi, L. J., Neri, R., Genzel, R., et al. 2013, ApJ, 768, 74 [§ 4.2.2, 12, A.2]  
 Thronson, Jr., H. A., & Telesco, C. M. 1986, ApJ, 311, 98 [§ 2.2.3]  
 Uzgil, B. D., Aguirre, J. E., Bradford, C. M., & Lidz, A. 2014, ApJ, 793, 116 [§ 1]  
 Visbal, E., & Loeb, A. 2010, J. Cosmology Astropart. Phys., 11, 16 [§ 1, 2, 2.1, 10, 3]  
 Walter, F., Decarli, R., Sargent, M., et al. 2014, ApJ, 782, 79 [§ 8, 4.3]  
 Wang, R., Carilli, C. L., Neri, R., et al. 2010, ApJ, 714, 699 [§ 3]  
 Watson, D., Christensen, L., Knudsen, K. K., et al. 2015, Nature, 519, 327 [§ 1]  
 Weiβ, A., Walter, F., & Scoville, N. Z. 2005, A&A, 438, 533 [§ 3]

## APPENDIX

### A. CO LUMINOSITY MODEL: SANITY CHECKS

#### A.1. Minimum CO-luminous Halo Mass

Here, we provide some justification for minimum CO-luminous halo mass of  $M_{\text{CO}, \text{min}} = 10^{10} M_{\odot}$  in our model. From a practical standpoint, this informs our choice of dark matter simulation, which is not complete below  $\sim 10^{10} M_{\odot}$ .

Physically, the cutoff is motivated by the idea that smaller galaxies at higher redshifts may not be CO-luminous. These galaxies may be less chemically evolved and/or less dusty than their low-redshift counterparts. As a result, CO in these systems might be less abundant and/or more likely to be dissociated by ionizing radiation. Such galaxies might not have significant CO emission despite active star formation.

A strict mass cutoff is chosen for simplicity and is unlikely to be the most realistic model. However, the effect of loosening it is well within other model uncertainties, so the cutoff we have chosen is acceptable for the scope of this paper.

To illustrate this quantitatively, Figure 11 shows the mean CO brightness temperature,  $\langle T_{\text{CO}} \rangle$ , as a function of this cut-off mass  $M_{\text{CO}, \text{min}}$ . Here,  $\langle T_{\text{CO}} \rangle$  has been obtained from the mean CO luminosity per volume, calculated as  $dL_{\text{CO}}/dV = \int_{M_{\text{CO}, \text{min}}}^{\infty} L_{\text{CO}}(M) (dn/dM) dM$  and converting to brightness temperature, where  $dn/dM$  is the analytic form of the halo mass function from Sheth et al. (2001). In our fiducial model at  $z = 2.4$ , lowering  $M_{\text{CO}, \text{min}}$  does increase the mean CO brightness

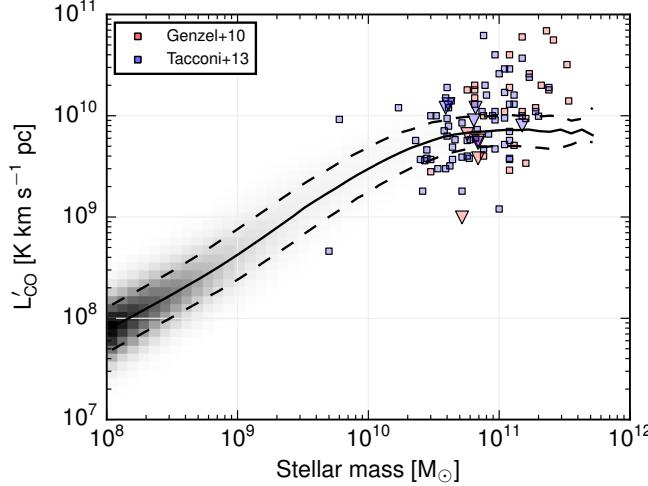


FIG. 12.— CO line luminosity vs. galaxy stellar mass. Stellar masses were calculated from the stellar mass–halo mass relation of Behroozi et al. (2013a), the same source for our SFR–halo mass relation. For comparison, the observations of Genzel et al. (2010) and Tacconi et al. (2013) have been plotted: squares are measured values, triangles are upper limits. Note that the observed  $L_{\text{CO}}$  data, in their respective studies, were converted to CO(1-0) luminosities from 3-2 and 2-1 lines, assuming a certain scaling between the different line luminosities. There is significant uncertainty in this scaling.

temperature, but with diminishing returns:  $L_{\text{CO}}(M)$  decreases more than the abundance of low-mass halos increases. To wit: in decreasing  $M_{\text{CO}, \min}$  by 1 dex from  $10^{11} M_{\odot}$  to  $10^{10} M_{\odot}$ ,  $\langle T_{\text{CO}} \rangle$  increases by  $\sim 25\%$ . In further decreasing  $M_{\text{CO}, \min}$  from  $10^{10} M_{\odot}$  to  $10^9 M_{\odot}$ ,  $\langle T_{\text{CO}} \rangle$  decreases by less than  $\sim 10\%$  — well within the uncertainty from  $SFR(M, z)$  alone — at  $M_{\min} = 10^{10} M_{\odot}$ .

Note that this particular justification is less valid if we assume  $L_{\text{CO}} \propto M$ , as in some previous models (§2.1 or Table 3). Incidentally, the fiducial models of those studies all assume  $M_{\text{CO}, \min} \leq 10^9 M_{\odot}$ , allowing more of the halo population to be CO-luminous. Note that these models do not take into account the rapid fall-off of the star formation efficiency below  $M^*$  that is indicated by several empirical studies (e.g. Behroozi et al. 2013a and references therein). For comparison, the same models from Figure 2 have also been plotted in Figure 11.

This justification is also more uncertain at higher redshifts. There, for a fixed halo mass, the halo mass function grows steeper with redshift, which may balance out the decrease in  $L_{\text{CO}}(M)$ .

### A.2. Comparison with Existing $L_{\text{CO}}$ -Mass Observations

Though sparse, measurements of  $L_{\text{CO}}$  exist for galaxies at these redshifts. Comparing the modeled  $L_{\text{CO}}(M)$  relation directly with observations is difficult, mainly because masses of dark matter halos are difficult to measure accurately. However, measurements of *stellar* mass are more commonly available from photometric data, and this is something we can compare.

Figure 12 shows CO luminosity vs. stellar mass, both as seen in observed galaxies and as predicted by our model. Observational data shown are from Tacconi et al. (2013) and Genzel et al. (2010) for  $z \sim 1$ –3 galaxies. Our model does not explicitly predict stellar mass, but we have calculated it using the stellar mass–halo mass relation from Behroozi et al. (2013a), which self-consistently yielded the SFR–halo mass relation used in our modeling.

The general consistency between model and data is reassuring, though not surprising, since the  $L_{\text{IR}}\text{--}L_{\text{CO}}$  relation that went into the model was empirically fit to existing observations. What bears the most emphasis is the near-total lack of data points for galaxies with low stellar mass ( $\lesssim 10^{10} M_{\odot}$ ), which, by number, comprise the overwhelming majority of the galaxy population. Note also that our model for the average properties includes quenched galaxies, while the data include only those galaxies detected in CO.

## B. IMPLEMENTATION DETAILS

For clarity, this section describes additional details for calculating CO brightness temperature and the power spectrum.

### B.1. Halo Luminosities to 3D Intensity Map

Once  $L_{\text{CO}}$  is calculated for all halos, they are binned on a grid according to their positions. This grid may be defined in observational units (RA–Dec–frequency) or comoving units. Converting between them requires us to specify a cosmology and rest-frame line frequency. In the two dimensions across the sky, we set the grid resolution to be 10 times finer than the beam size, i.e. the voxel length in the RA and Dec directions was  $\theta_{\text{FWHM}}/10$ .

On a comoving grid, the observed brightness temperature from a voxel of comoving volume  $V_{\text{vox}}$ , containing total luminosity  $L_{\text{CO, vox}}$ , emitted at rest frequency  $\nu_{\text{rest}}$ , from redshift  $z$  is

$$\begin{aligned} T_{\text{CO, vox}} &= \frac{c^3(1+z)^2}{8\pi k_B \nu_{\text{rest}}^3 H(z)} \frac{L_{\text{CO, vox}}}{V_{\text{vox}}} \\ &= 3.1 \times 10^4 \mu\text{K} (1+z)^2 \left( \frac{\nu_{\text{rest}}}{\text{GHz}} \right)^{-3} \left( \frac{H(z)}{\text{km s}^{-1} \text{Mpc}^{-1}} \right)^{-1} \left( \frac{L_{\text{CO, vox}}}{L_{\odot}} \right) \left( \frac{V_{\text{vox}}}{\text{Mpc}^3} \right)^{-1}. \end{aligned}$$

### B.2. Fourier convention and power spectrum

We calculate the power spectrum through the Fourier transform,  $F(\mathbf{k})$ , of the 3D CO temperature cube. The power spectrum  $P(\mathbf{k})$  (specifically, power spectral density) is then

$$P(\mathbf{k}) = V_{\text{surv}}^{-1} |F(\mathbf{k})|^2 \quad (\text{B1})$$

where  $V_{\text{surv}}$  is the comoving volume of the survey.

The Fourier transform requires a choice of convention, which does not matter as long as it is consistently applied. However, because conventions and numerical implementations may differ slightly, we state ours here.

For the continuous Fourier transform  $F(\mathbf{k})$ , we use the following (non-unitary angular frequency) convention:

$$\text{Forward: } F(\mathbf{k}) = \int f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d^3\mathbf{x}$$

with the reverse transform being  $f(\mathbf{x}) = (2\pi)^{-3} \int F(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} d^3\mathbf{k}$ . We only need the forward transform for  $P(k)$ , so this choice conveniently avoids additional factors of  $2\pi$ . Our discrete Fourier transform  $F_{\mathbf{k}}$  is defined as

$$F_{\mathbf{k}} = \sum_n^{N-1} a_n \exp \left[ -2\pi i \frac{nk}{N} \right], \quad k = 0, \dots, n-1 \quad (\text{B2})$$

and is thus related to the continuous Fourier transform  $F(\mathbf{k})$  as  $F(\mathbf{k}) \approx \Delta x \Delta y \Delta z F_{\mathbf{k}}$ , where  $\Delta x$ ,  $\Delta y$ , and  $\Delta z$  are the comoving voxel lengths.

### B.3. MCMC Implementation

We used the *emcee* Python package (v2.1.0, [Foreman-Mackey et al. 2013](#)), an implementation of an affine-invariant ensemble algorithm ([Goodman & Weare 2010](#)), to perform the MCMC analysis. This approach prioritizes ease of implementation over computational speed, which we believe to be justified for the scope of this work. To sample the five-dimensional parameter space, we used an ensemble of 200 walkers taking 2500 steps (500,000 samples in total), after a fairly generous burn-in phase of 500 steps.

At each sampled position in parameter space, we calculate a power spectrum  $P_{\text{model}}(k)$  by generating a new intensity map from one of 100 realizations of the survey volume. Because of the finite volume of these realizations,  $P_{\text{model}}$  is not immune to sample variance, so it carries its own error bars,  $\sigma_{P_{\text{model}}}$ .

The "observed" intensity map, in reality, would appear smoothed by the telescope beam (see Figure 1), i.e. convolved with a 2D Gaussian filter in each frequency channel. This attenuates the apparent power spectrum, especially at scales smaller than the beam. In practice, the MCMC procedure would require this convolution to be repeatedly computed, which is moderately expensive even with Fourier transforms. Instead, we calculate the "unsmoothed" power spectrum for  $P_{\text{obs}}$  and  $P_{\text{model}}$  but increase the error bars on both by a factor  $[W(k)]^{-1}$  (see Eq. C4). Thus, the resolution limits are encoded in the error bars ( $\sigma_{P_{\text{model}}}$  and  $\sigma_{P_{\text{obs}}}$ ), rather than in  $P_{\text{model}}$  and  $P_{\text{obs}}$ .

Within additive constants, the log-likelihood expression is

$$\ln \mathcal{L} = -\frac{1}{2} \sum_k \left\{ \frac{[P_{\text{model}}(k) - P_{\text{obs}}(k)]^2}{\sigma_{P_{\text{model}}}(k) + \sigma_{P_{\text{obs}}}(k)} + \ln [\sigma_{P_{\text{model}}}(k) + \sigma_{P_{\text{obs}}}(k)] \right\} \quad (\text{B3})$$

where  $\sigma_{P_{\text{obs}}}$  is given by Eq. C1, and  $\sigma_{P_{\text{model}}} = P_{\text{model}} / \sqrt{N_{\text{modes}}} / W$ , where  $N_{\text{modes}}$  and  $W$  are functions of  $k$  given by Eqs. C3 and C4, respectively.

The 3D Fourier transform makes calculating the power spectrum the most computationally expensive part of each MCMC step, particularly as dynamic range increases (larger survey volumes and finer resolution).

### C. ERROR BARS ON THE POWER SPECTRUM

In this section, we summarize the calculation of  $\sigma_P(k)$ , the uncertainty in the spherically averaged power spectrum, or the error bars on  $P(k)$ —and by extension,  $\Delta^2(k)$ . This is essential for quantifying how precisely an intensity mapping experiment can detect a signal, as well as cosmological structure within that signal.

The details that follow are previously mentioned in, e.g., [Lidz et al. \(2011\)](#) and [Gong et al. \(2012\)](#). We find that a small but non-negligible correction in the calculation is necessary (noted below). Otherwise, we have mainly redescribed the procedure here for completeness and clarity.

This calculation accounts for three sources of uncertainty in the power spectrum:

1. Sample variance.
2. Thermal noise variance, from the instrument.
3. Limited resolution.

In a real experiment, there can certainly be additional systematics to consider, but this calculation is an appropriate starting point. The equation that summarizes the calculation is

$$\sigma_P(k) = \left( \underbrace{\frac{P(k)}{\sqrt{N_{\text{modes}}(k)}}}_{\text{Sample variance}} + \underbrace{\frac{P_n(k)}{\sqrt{N_{\text{modes}}(k)}}}_{\text{Thermal noise}} \right) \underbrace{\frac{1}{W(k)}}_{\text{Resolution}} \quad (\text{C1})$$

where

- $P(k)$  is the (observed) spherically averaged power spectrum, as a function of  $k$
- $P_n(k)$  is the thermal noise power spectrum:

$$P_n(k) = \sigma_n^2 V_{\text{vox}} \quad (\text{C2})$$

- $N_{\text{modes}}(k)$  is the number of measured modes at  $k$  in a bin of width  $\Delta k$ :

$$N_{\text{modes}}(k) = \frac{k^2 \Delta k V_{\text{surv}}}{4\pi^2} \quad (\text{C3})$$

- $W(k)$  is a factor that attenuates the power spectrum, arising from limited spatial resolution and mainly affecting high  $k$  beyond resolvable scales:

$$W(k) = e^{-k^2 \sigma_{\perp}^2} \int_0^1 e^{-k^2 (\sigma_{\parallel}^2 - \sigma_{\perp}^2) \mu^2} d\mu \quad (\text{C4})$$

We explain Eqs. C2, C3, and C4 in the subsections below. Before we do, the following points may help to clarify notation, language, and assumptions:

- Redshift is  $z$ , at which  $H(z)$  is the Hubble parameter,  $R(z)$  is the comoving radial distance, and  $D_L(z)$  is the luminosity distance. We assume a single redshift  $1+z = \nu_{\text{rest}}/\nu_{\text{obs}} = \lambda_{\text{obs}}/\lambda_{\text{rest}}$  for the volume, an acceptable approximation for the surveys being considered.
- Comoving Cartesian coordinates are labeled  $(x_1, x_2, x_3)$  in order to avoid confusion with redshift.
- $x_1$  and  $x_2$  are perpendicular to the line of sight (denoted  $\perp$ , mapping to RA/Dec).  $x_3$  is parallel to the line of sight (denoted  $\parallel$ , mapping to redshift or observed frequency).
- The instrument has system temperature  $T_{\text{sys}}$  and number of observing feeds  $N_{\text{feeds}}$ . The full bandwidth of the spectrometer is  $\Delta\nu$ , centered on an observing frequency  $\nu_{\text{obs}}$  and divided into frequency channels of width  $\delta\nu$ . The observing beam has a Gaussian profile, with full-width half-maximum angle  $\theta_{\text{FWHM}}$  ( $\sigma_{\text{beam}} = \theta_{\text{FWHM}}/\sqrt{8\ln 2}$ ). The total observing time on the survey area is  $\tau_{\text{obs}}$ .
- The smallest resolvable 3D volume element is approximated as a cubic “voxel,” while the sky projection of all voxels along a line of sight is a “pixel” (subscripted “vox” and “pix”).  $V_{\text{vox}}$  is the comoving volume of a voxel, while  $\Omega_{\text{pix}}$  is the solid angle on the sky subtended by a pixel. The expressions for each are

$$\Omega_{\text{pix}} = \sigma_{\text{beam}}^2 \quad \text{and} \quad V_{\text{vox}} = [R(z)\sigma_{\text{beam}}]^2 \left[ \frac{c}{H(z)} \frac{\delta\nu(1+z)^2}{\nu_{\text{rest}}} \right]. \quad (\text{C5})$$

- $V_{\text{surv}} \approx L_1 L_2 L_3$  is the comoving volume of the survey, while  $\Omega_{\text{surv}}$  is the solid-angle sky coverage of the survey.

### C.1. Noise Power Spectrum [ $P_n(k)$ , Eq. C2]

Consider any voxel in the observed 3D temperature cube, covering a sky area  $\Omega_{\text{pix}}$  and frequency bin  $\delta\nu$  as previously described. We assume that the random thermal noise fluctuations, i.e. a white Gaussian noise spectrum. By the radiometer equation, the rms temperature fluctuation per voxel is

$$\sigma_n = \frac{T_{\text{sys}}}{\sqrt{N_{\text{feeds}} \delta\nu \tau_{\text{pix}}}} \quad (\text{C6})$$

where  $\tau_{\text{pix}} = \tau_{\text{obs}}(\Omega_{\text{pix}}/\Omega_{\text{surv}})$  is the observing time per sky pixel.

The averaged power spectrum of this noise is a constant:

$$P_n(k) = \sigma_n^2 V_{\text{vox}}.$$

### C.2. Number of Modes [ $N_{\text{modes}}(k)$ , Eq. C3]

Because  $P(k)$  is obtained from averaging  $P(\mathbf{k})$  over a finite number of modes (points in  $k$ -space), we need to know the number of independently measured modes used to compute the average. We label this number  $N_{\text{modes}}$  and thus a factor  $1/\sqrt{N_{\text{modes}}}$  ultimately enters into the calculation of  $\sigma_P(k)$ .

Written explicitly,  $N_{\text{modes}}$  depends on both the scale of the fluctuation,  $k$ , and the choice of bin width,  $\Delta k$ . Consider a radial  $k$ -space bin (i.e. spherical shell) between  $k$  and  $k + \Delta k$ . Assuming  $\Delta k$  is chosen, the number of modes within this shell is

$$N_{\text{modes}}(k) = n(k) \cdot 4\pi k^2 \Delta k \quad (\text{C7})$$

where  $n(k)$  is the number density of modes in  $k$ -space. Recalling the similar “density of states” calculation in statistical mechanics,  $n(k)$  is

$$n(k) = \left( \frac{2\pi}{L_1} \frac{2\pi}{L_2} \frac{2\pi}{L_3} \right)^{-1} = \frac{V_{\text{surv}}}{8\pi^3}. \quad (\text{C8})$$

However, the power spectrum of a real-valued function is symmetric about  $\mathbf{k} = 0$ , so only half of the modes contain independent information. This introduces a factor  $1/2$  into the final expression, which is (from combining Eq. C7 and C8):

$$N_{\text{modes}}(k, \Delta k) = \frac{k^2 \Delta k V_{\text{surv}}}{4\pi^2}.$$

### C.3. Resolution Limits [ $W(k)$ , Eq. C4]

Finite spatial resolution means that information from high- $k$  modes is lost, due to smoothing of features smaller than the beam or channel width. The basic reasoning here is that if the smoothed power spectrum  $P_{\text{sm}}(k)$  results from attenuating the “intrinsic” power spectrum  $P(k)$  by a factor  $W(k)$  after smoothing, the uncertainty should scale as  $\sigma_P(k) \propto P(k)/P_{\text{sm}}(k) = 1/W(k)$ .

With arbitrarily fine resolution, the temperature field is  $T(x_1, x_2, x_3)$ , which has a power spectrum  $P(k)$ . In reality, we measure a smoothed field,  $T_{\text{sm}}(x_1, x_2, x_3)$ , which is the original field convolved with a Gaussian in each direction:

$$T_{\text{sm}}(\mathbf{x}) = T(x_1, x_2, x_3) * G(x_1|\sigma_1) * G(x_2|\sigma_2) * G(x_3|\sigma_3) \quad (\text{C9})$$

where  $G(x_i|\sigma_i) = (2\pi\sigma_i)^{-1/2} \exp[-x_i^2/(2\sigma_i^2)]$  (similarly for  $x_2$  and  $x_3$ ) is a normalized Gaussian function, and “ $*$ ” indicates convolution. In reality, the beam profile may not be perfectly Gaussian, and the frequency channels would be discrete bins. However, the approximation will suffice here.

Angular resolution determines  $\sigma_1$  and  $\sigma_2$ , while frequency resolution determines  $\sigma_3$ . Specifically,

$$\sigma_1 = \sigma_2 \equiv \sigma_{\perp} = R(z)\sigma_{\text{beam}} \quad \text{and} \quad \sigma_3 \equiv \sigma_{\parallel} = \frac{c}{H(z)} \frac{\delta_{\nu}(1+z)}{\nu_{\text{obs}}}. \quad (\text{C10})$$

Since a convolution in real space is simple multiplication in Fourier space, the Fourier transform and power spectrum are

$$\tilde{T}_{\text{sm}}(\mathbf{k}) = \tilde{T}(k_1, k_2, k_3) e^{-(k_1^2/\xi_1^2 + k_2^2/\xi_2^2 + k_3^2/\xi_3^2)/2} \quad \text{and} \quad P_{\text{sm}}(\mathbf{k}) = P(k_1, k_2, k_3) e^{-(k_1^2/\xi_1^2 + k_2^2/\xi_2^2 + k_3^2/\xi_3^2)} \quad (\text{C11})$$

where  $\xi_1 = \xi_2 = 1/\sigma_{\perp}$  and  $\xi_3 = 1/\sigma_{\parallel}$ . The Fourier convention used here ensures the transformed Gaussian carries no normalization constant.

We note a small but consequential correction: Lidz et al. (2011) and some following studies state  $\xi_{\perp} = 2\pi/[R(z)\sigma_{\text{beam}}]$ , but we find from the calculation above that it should actually be  $\xi_{\perp} = 1/\sigma_{\perp} = 1/[R(z)\sigma_{\text{beam}}]$  (without the factor of  $2\pi$ ).

To get the spherically averaged power spectrum, define the variable  $\mu = \cos \vartheta$ , where  $\vartheta$  is the spherical polar angle in  $k$ -space. Then we define components of  $k$  parallel and perpendicular to the  $z$  direction,  $k_{\parallel}$  and  $k_{\perp}$ , as

$$k_{\parallel} = k_3 = k\mu \quad \text{and} \quad k_{\perp} = \sqrt{k^2 - k_{\parallel}^2} = k\sqrt{1 - \mu^2} \quad (\text{C12})$$

and the expression for the power spectrum becomes

$$\begin{aligned} P_{\text{sm}}(\mathbf{k}) &= P(k_{\perp}, k_{\parallel}) e^{-(\sigma_{\perp}^2 k_{\perp}^2 + \sigma_{\parallel}^2 k_{\parallel}^2)} \\ &= P(k, \mu) e^{-k^2 \sigma_{\perp}^2} e^{-\mu^2 k^2 (\sigma_{\parallel}^2 - \sigma_{\perp}^2)}. \end{aligned} \quad (\text{C13})$$

To get the spherically averaged power spectrum, this needs to be averaged over  $0 < \mu < 1$  (the upper half-space, since only half the modes are independent) at fixed  $k$ . In this paper, we assume  $P(k, \mu)$  is isotropic, so only the exponential needs to be averaged:

$$P_{\text{sm}}(k) = P(k) \underbrace{\int_0^1 e^{-k^2 (\sigma_{\parallel}^2 - \sigma_{\perp}^2) \mu^2} d\mu}_{W(k)}. \quad (\text{C14})$$

The expression after  $P(k)$  is  $W(k)$ , which can be integrated numerically.

#### D. OPTIMIZING OBSERVING PARAMETERS

The expression for  $\sigma_P(k)$  allows us to draw some conclusions about the effect of varying survey and instrument parameters. See [Breysse et al. \(2014\)](#) for a similar analysis with  $C_\ell$  coefficients, i.e. a 2D power spectrum.

We ignore the factor  $W(k)$  for this analysis, since its main effect is to limit the maximum  $k$  (smallest scale) at which meaningful measurements can be made. Certainly, it will also slightly suppress power at low  $k$  (large scales), but this should not be a dramatic effect. Then, the expression for  $\sigma_P(k)$ , showing explicit dependences on instrument and survey parameters, is ultimately

$$\sigma_P(k) = \frac{2\pi}{k\sqrt{\Delta k \Delta\nu}} \left[ \frac{1}{\sqrt{\lambda_{\text{rest}} \Omega_{\text{surv}}}} \frac{\sqrt{H(z)}}{D_L(z)} P(k) + \frac{T_{\text{sys}}^2}{N_{\text{feeds}} \tau_{\text{obs}}} \sqrt{\lambda_{\text{rest}} \Omega_{\text{surv}}} \frac{D_L(z)}{\sqrt{H(z)}} \right]. \quad (\text{D1})$$

where  $D_L(z) = (1+z)R(z)$  is the luminosity distance. Note that redshift  $z$  is directly determined by  $\nu_{\text{obs}}$ .

Written in this form, it is easier to see how instrument and survey parameters ultimately affect sensitivity. At a given scale  $k$ , the following adjustments will decrease  $\sigma_P(k)$ :

- **Increasing bandwidth:**  $\sigma_P \propto 1/\sqrt{\Delta\nu}$ . This expands the survey volume, increasing the density of modes in the  $k_{\parallel}$  direction and thus  $N_{\text{modes}}$ . However, the observed redshift range will eventually grow large enough that cosmic evolution across the volume becomes important.
- **Decrease instrument noise:** if  $\sigma_P$  is noise-dominated,  $\sigma_P \propto T_{\text{sys}}^2/N_{\text{feeds}}\tau_{\text{obs}}$ . In that case, decreasing  $T_{\text{sys}}$ , increasing  $N_{\text{feeds}}$ , or increasing  $\tau_{\text{obs}}$  will all serve to increase sensitivity. Eventually, though, decreasing system noise yields diminishing returns as sample variance becomes the dominant component of  $\sigma_P$ .
- **Optimize survey area.** For a given mode  $k$ , the optimal survey area results in equal sample variance ( $\propto \Omega_{\text{surv}}^{-1/2}$ ) and noise variance ( $\propto \Omega_{\text{surv}}^{1/2}$ ). In steradians, this is

$$\Omega_{\text{surv, opt}} = \frac{1}{\lambda_{\text{rest}}} \underbrace{\frac{N_{\text{feeds}}\tau_{\text{obs}}}{T_{\text{sys}}^2}}_{\text{From noise}} \underbrace{\frac{H(z)}{[D_L(z)]^2}}_{\text{From } \nu_{\text{obs}}} P(k). \quad (\text{D2})$$

Because  $\Omega_{\text{surv, opt}} \propto P(k)$ , the predicted optimal area is model-dependent.

As noted in §3.1, we can maximize the total detection significance,  $\text{SNR}_{\text{tot}}$  (Eq. 7), by optimizing the area. For the observing parameters in this study, we obtain 0.6 and 9.2 deg<sup>2</sup> for the pathfinder and full experiments, respectively. Certainly, these optimal areas differ from our adopted survey areas (2.5 and 6.25 deg<sup>2</sup>), but the improvement in  $\text{SNR}_{\text{tot}}$  is not dramatic (7.89 to 9.48 for pathfinder; 144 to 146 for full).

However, in practice, there are other factors to consider. The minimum survey area cannot be arbitrarily small, because the arrangement of multiple feed horns on the instrument images a pattern of multiple beams on the sky. This pattern must be scanned across the sky to adequately cover the survey area, and limitations on the scanning strategy effectively set the minimum size of the survey area. This motivated our estimate of 2.5 and 6.25 deg<sup>2</sup> as the survey areas for the pathfinder and full experiments, respectively.

The maximum survey area can be as large as the entire sky, but increasing total area would reduce the amount of integration time *per unit area*. This can decrease sensitivity to high- $k$  (shot noise) modes, which contain important astrophysical information.

It is worth noting that the following parameters do *not*, for the simple assumptions above, affect  $\sigma_P$  (neglecting secondary effects such as redshift evolution across the volume and the curvature of the boundary surface, which we expect to be subdominant for the volumes being considered):

- **Resolution.** Assuming the spatial scales being probed are much larger than the resolution limit, refining the resolution does not decrease  $\sigma_P$  significantly.
- **Survey area aspect ratio.** For a fixed rectangular survey area, its relative dimensions have no effect. As the length grows, the width shrinks. In  $k$ -space, the density of modes becomes sparser in one direction but greater in the other, and  $N_{\text{modes}}$  does not change.

#### E. OBTAINING DATA FROM OBRESCHKOW ET AL. (2009)

For the results of §3.3.1, we downloaded the lightcone data of [Obreschkow et al. \(2009b\)](#) from their publicly accessible database. We selected all galaxies within a central 6.25 deg<sup>2</sup> area, located at comoving distances between 5770 and 6285 Mpc, which encloses the redshift range  $2.4 < z < 2.8$  in our assumed cosmology. From each simulated galaxy, we obtained four quantities: RA, Dec, comoving distance, and CO(1-0) flux. On the query form page<sup>12</sup>, the exact SQL query given was

<sup>12</sup> [http://s-cubed.physics.ox.ac.uk/queries/new?  
sim=s3\\_sax](http://s-cubed.physics.ox.ac.uk/queries/new?sim=s3_sax)

```
select ra, decl, cointflux_1, distance
from galaxies_line
where ra between -1.25 and 1.25
      and decl between -1.25 and 1.25
      and distance between 5770 and 6285
```

Since CO(1-0) flux was provided as an observed velocity-integrated flux,  $S^V$  (units of Jy km s $^{-1}$ ), we converted to an intrinsic brightness temperature luminosity,  $L^T$  (units of K km s $^{-1}$  pc $^2$ ) following the derived equation in the appendix of [Obreschkow et al. \(2009a\)](#):

$$\frac{L^T}{\text{K km s}^{-1} \text{ pc}^2} = 3.255 \times 10^7 \left( \frac{\nu_{\text{obs}}}{\text{GHz}} \right)^{-2} \left( \frac{D_L}{\text{Mpc}} \right)^2 (1+z)^{-3} \frac{S^V}{\text{Jy km s}^{-1}} \quad (\text{E1})$$

From these galaxy luminosities, we generated CO intensity maps and power spectra as described in §[2.2.5](#).

TABLE 3  
PREVIOUS CO INTENSITY MAPPING MODELS

Paper	Redshifts	CO luminosity model	
Righi et al. (2008)	$z \sim 0 - 10$	<b>Calculating SFR:</b> SFR assumed to be driven by major mergers. Stellar mass $M^*$ formed in a merger of total mass $M = M_1 + M_2$ calculated as: $M^* = 4 \frac{\Omega_b}{\Omega_m} \eta \frac{M_1 M_2}{M} = 3.5 \times 10^{-2} \frac{M_1 M_2}{M}$ Merger rate was calculated from extended Press-Schechter formalism (Lacey & Cole 1993).	<b>Converting SFR to <math>L_{\text{CO}}</math>:</b> Based on M82 observations (Weiß et al. 2005): $L_{\text{CO}} = 3.7 \times 10^3 \text{ SFR}$
Visbal & Loeb (2010)	$z \gtrsim 3$	<b>Calculating SFR:</b> SFR proportional to halo mass, $M_{\text{halo}}$ : $\text{SFR} = 6.2 \times 10^{-11} \left( \frac{1+z}{3.5} \right)^{3/2} M_{\text{halo}}$	<b>Converting SFR to <math>L_{\text{CO}}</math>:</b> Based on M82 observations (Weiß et al. 2005): $L_{\text{CO}} = 3.7 \times 10^3 \text{ SFR}$
Gong et al. (2011)	$z \sim 6 - 8$	<b>Calculating <math>L_{\text{CO}}</math>:</b> Fit a relation of the form $L_{\text{CO}} = L_0 \left( \frac{M_{\text{halo}}}{M_c} \right)^b \left( 1 + \frac{M_{\text{halo}}}{M_c} \right)^{-d}$ to the results from semi-analytic modeling by Obreschkow et al. (2009a).	At $z = 6, 7$ , and $8$ respectively: $L_0 = 4.3 \times 10^6, 6.2 \times 10^6, 4.0 \times 10^6 \text{ L}_{\odot}$ $b = 2.4, 2.6, 2.8$ $M_c = 3.5 \times 10^{11}, 3.0 \times 10^{11}, 2.0 \times 10^{11} \text{ M}_{\odot}$ $d = 2.8, 3.4, 3.3$
Carilli (2011)	$z \sim 6 - 10$	<b>Calculating SFR:</b> Cosmic SFR density calculated as the value required to reionize the universe at a given redshift (Bunker et al. 2010; Madau et al. 1999).	<b>Converting SFR to <math>L_{\text{CO}}</math>:</b> Combining SFR- $L_{\text{IR}}$ (Kennicutt 1998) and disk galaxy $L_{\text{IR}}-L_{\text{CO}}$ (Daddi et al. 2010) relations: $L_{\text{CO}} = 1.0 \times 10^4 \text{ SFR}$
Lidz et al. (2011)	$z \gtrsim 6$	<b>Calculating SFR:</b> SFR proportional to halo mass, $M_{\text{halo}}$ : $\text{SFR} = 1.7 \times 10^{-10} M_{\text{halo}}$	<b>Converting SFR to <math>L_{\text{CO}}</math>:</b> Combining SFR- $L_{\text{IR}}$ (Kennicutt 1998) and $L_{\text{IR}}-L_{\text{CO}}$ (Wang et al. 2010) relations: $L_{\text{CO}} = 3.2 \times 10^4 \text{ SFR}$ Note: as originally derived, $L_{\text{CO}} \propto (\text{SFR})^{3/5}$ . Linear scaling adopted for simplicity.
Pullen et al. (2013)	$z \sim 2 - 4$	<b>Calculating SFR:</b> Model A: SFR proportional to $(M_{\text{halo}})^{5/3}$ : $\text{SFR} = 1.2 \times 10^{-11} (M_{\text{halo}})^{5/3}$ Model B: SFR calculated from empirically fit SFR Schechter functions. (Smit et al. 2012)	<b>Converting SFR to <math>L_{\text{CO}}</math>:</b> Same as Lidz et al. (2011): $L_{\text{CO}} = 3.2 \times 10^4 (\text{SFR})^{3/5}$ Note: for Model A, the effective model is still $L_{\text{CO}} \propto M_{\text{halo}}$ . For Model B, power spectra were obtained by multiplying those of A by a factor $\langle T_{\text{CO,B}} \rangle^2 / \langle T_{\text{CO,A}} \rangle^2$ ( $\approx 4.8^2$ at $z = 3$ ).

NOTE. — For consistent comparison, relations here were simplified using the authors' given fiducial assumptions, but we refer the reader to the original papers for full parameterizations and details. See Figure 2 for a comparison of  $L_{\text{CO}}(M_{\text{halo}})$  between some of these models.